

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 22, 2001, 00:18:34, Search time 1471.74 seconds
(without alignments)
6659,142 Million cell updates/sec

```
Title: US-09-463-480-3
Perfect score: 625
Sequence: 1 qcatccatcacacgaagg...aaataaaaataaaaataaaa
```

Scoring table: IDENTITY_NJC
Gapop 10.0 Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

Database :
GenEmbl.*
1: qb_bal.*
2: qb_ba2.*
3: qb_ba3.*
4: qb_in1.*
5: qb_in2.*
6: qb_in3.*
7: qb_om.*
8: qb_ov.*
9: qb_par1.*
10: qb_pat2.*
11: qb_ph.*
12: qb_p11.*
13: qb_p12.*
14: qb_p13.*
15: qb_p14.*
16: em_bal.*
17: em_ba2.*
18: em_fun.*
19: em_hqo_hum.*
20: em_hqo_inv.*
21: em_hqo_rod.*
22: em_hq_hum1.*
23: em_hq_hum2.*
24: em_hq_hum3.*
25: em_hq_hum4.*
26: em_hq_hum5.*
27: em_hq_hum6.*
28: em_hq_hum7.*
29: em_hq_hum8.*
30: em_hq_invl.*
31: em_hq_invl2.*
32: em_hq_other.*
33: em_hq_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

```

44:	em_ov:
45:	em_pat:
46:	em_ph:
47:	em_pl:
48:	em_to:
49:	em_sts:
50:	em_sy:
51:	em_un:
52:	em_vl:
53:	dt_stsl:
54:	st_stsl:
55:	st_sts:
56:	st_sy:
57:	st_un:
58:	ql_vl1:
59:	ql_vl2:
60:	ql_vl3:
61:	ql_vl4:
62:	ql_vl5:
63:	ql_vl6:
64:	ql_vl7:
65:	ql_vl8:
66:	ql_vl9:
67:	ql_vl10:
68:	ql_vl11:
69:	ql_vl12:
70:	ql_vl13:
71:	ql_vl14:
72:	ql_vl15:
73:	ql_vl16:
74:	ql_vl17:
75:	ql_vl18:
76:	ql_vl19:
77:	ql_vl20:
78:	ql_vl21:
79:	ql_vl22:
80:	ql_vl23:
81:	ql_vl24:
82:	ql_vl25:
83:	ql_vl26:
84:	ql_vl27:
85:	ql_vl28:
86:	ql_vl29:
87:	ql_vl30:
88:	ql_vl31:
89:	ql_vl32:
90:	ql_vl33:
91:	ql_vl34:
92:	ql_vl35:
93:	ql_vl36:
94:	ql_vl37:
95:	ql_vl38:
96:	ql_vl39:
97:	ql_vl40:
98:	ql_vl41:
99:	ql_vl42:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DR	ID	Description
1	62.5	100.0	625	12	AF110779	AF110779 Lillium fo
2	66.8	10.7	17839	4	AE001384	AE001384 Plasmodi
3	65.2	10.4	750	93	U85M02288	U85M02288 Homo sapi
4	65.2	10.4	321063	84	U85M044P3	U85M044P3 Plasmodi
5	63.8	10.2	865	5	U85M05V71	U85M05V71 17 end of
6	63.8	10.2	2710	6	U85M05V71	U85M05V71 17 end of
7	62	9.9	924	53	U85M07A5L	U85M07A5L 17 end of
8	61.4	9.8	2976	6	U85M09478	U85M09478 Dictyosteli



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 22, 2001, 00:23:00, Search time 52.75 seconds
(without alignments)
1248.654 Million cell updates/sec

Title: US-09-463-480-3

Perfect score: 625

Sequence: 1 accatccatcacacagaag.....aaaaaaaaaaaaaaaaaaaaaa 625

Scoring table: IDENTITY_NUC

Gapop 10 0 . Gapext 1 0

Searched: 317530 seqs, 92630169 residues

Total number of hits satisfying chosen parameters. 635060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /com2_6/ptdata/1/ina/5A.COMB.seq.*
- 2: /com2_6/ptdata/1/ina/5B.COMB.seq.*
- 3: /com2_6/ptdata/1/ina/6A.COMB.seq.*
- 4: /com2_6/ptdata/1/ina/6B.COMB.seq.*
- 5: /com2_6/ptdata/1/ina/6C.COMB.seq.*
- 6: /com2_6/ptdata/1/ina/6D.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58.6	9.4	857	1	US-08-308-883-1
2	58.6	9.4	857	1	US-08-730-163-1
3	58.6	9.4	857	4	US-08-256-799-1
4	58.6	9.4	857	4	US-08-462-437-1
5	58.2	9.3	3138	1	US-07-867-106-4
6	56.8	9.1	991	3	US-08-924-747-25
7	56.8	9.1	991	4	US-09-247-273R-25
8	56.8	9.1	991	4	US-09-296-715-25
9	56.6	9.1	1474	4	US-08-821-994-64
10	56.2	9.0	703	4	US-09-413-309-6
11	56.2	9.0	1582	3	US-08-545-166R-10
12	56.2	9.0	1582	3	US-08-545-166B-12
13	56	9.0	1172	1	US-07-945-288-9
14	56	9.0	1172	1	US-08-462-831-9
15	56	9.0	1172	1	US-08-461-809-9
16	56	9.0	1172	1	US-08-461-441-9
17	56	9.0	1172	5	PCT US93/08518-9
18	55.6	8.9	5852	1	US-07-867-106-2
19	55.4	8.9	1578	4	US-09-416-050A-1
20	55.4	8.9	1578	4	US-09-664-800-1
21	55.4	8.9	1578	4	US-09-665-309-1
22	55.2	8.8	3527	2	US-08-905-9552-7
23	54.8	8.8	1534	1	US-08-300-903A-6
24	54.6	8.7	1075	4	US-08-400-006R-6
25	54.2	8.7	2852	3	US-09-027-137-2
26	53.8	8.6	1364	1	US-08-265-087-3
27	53.8	8.6	1364	1	US-08-621-493-3

```

28 53.8 8.6 1364 2 US-08-965-688-3 Sequence 3, Appli
29 53.8 8.6 1364 4 US-09-260-173-3 Sequence 3, Appli
30 10660 2 US-08-267-803B-8 Sequence 8, Appli
31 10660 4 US-09-041-886-16 Sequence 16, Appli
32 53.4 8.5 1046 1 US-08-361-467B-4 Sequence 4, Appli
33 53.4 8.5 1046 1 US-08-484-342R-4 Sequence 4, Appli
34 53.2 8.5 1641 1 US-08-300-903A-8 Sequence 8, Appli
35 53.2 8.5 1776 3 US-06-655-392-10 Sequence 10, Appli
36 53.2 8.5 5173 1 US-08-242-677-1 Sequence 1, Appli
37 53 8.5 8920 2 US-08-446-855A-1 Sequence 1, Appli
38 53 8.5 8920 4 US-09-150-741-1 Sequence 1, Appli
39 52.8 8.4 1700 2 US-08-897-340-4 Sequence 4, Appli
40 52.8 8.4 1700 4 US-09-252-229-4 Sequence 4, Appli
41 52.8 8.4 3238 5 PCT US94/10880-5 Sequence 5, Appli
42 52.6 8.4 2447 2 US-09-014-969-14 Sequence 14, Appli
43 52.4 8.4 2550 6 5258287-23 Patent No. 5258287
44 52.4 8.4 2887 5 PCT US96/10521-14 Sequence 14, Appli
45 52.2 8.4 1098 4 US-09-248-335-35 Sequence 35, Appli

```

ALIGNMENTS

```

RESULT 1
US-08-308-883-1
: Sequence 1, Application US/08408884
: Patent No. 5576400
: GENERAL INFORMATION:
: APPLICANT: Mukerji, P. A.
: APPLICANT: Prieto, P. A.
: APPLICANT: Seo, A. E.-Y.
: APPLICANT: Baxter, J. H.
: APPLICANT: Cummings, R.D.
: TITLE OF INVENTION: Method for inhibition of human rotavirus infection.
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Louie R. Drayer
: ADDRESSEE: KES Products Division
: ADDRESSEE: Abbott Laboratories
: STREET: 625 Cleveland Avenue
: CITY: Columbus
: STATE: Ohio
: COUNTRY: United States
: ZIP: 43215
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb storage
: COMPUTER: Apple Macintosh
: OPERATING SYSTEM: Macintosh System 7.1
: SOFTWARE: ClarisWorks 1.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER US/08408883
: FILING DATE: 16-Sep-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA: No. 5576400 applicable
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (614) 624-3774
: TELEFAX: (614) 624-3074
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 857 base pairs
: TYPE: Nucleic acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: MOLECULE TYPE: cDNA
: DESCRIPTION: Human milk kappa casein
: HYPOTHETICAL: No
: ANTI-SENSE:
: FRAGMENT TYPE:
: ORIGINAL SOURCE: Human
: ORGANISM: homo sapiens
: STRAIN:
: INDIVIDUAL ISOLATE:

```


Db 839 A 839

```
RESULT 3
US-08-456-799-1
: Sequence 1, Application US/08256799
: Patent No. 6232094
: GENERAL INFORMATION:
: APPLICANT: HANSSON, Lennart
: APPLICANT: STROEMQVIST, Mats
: APPLICANT: BERGSTROEM, Sven
: APPLICANT: HERNELL, Olie
: APPLICANT: Toernell, Jan
: TITLE OF INVENTION: DNA ENCODING KAPPA-CASEIN, PROCESS FOR
: TITLE OF INVENTION: OBTAINING THE PROTEIN AND USE THEREOF
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWDY AND NEIMARK
: STREET: 419 Seventh Street, N W , Suite 300
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/256,799
: FILING DATE: 06-DEC-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DK 89/92
: FILING DATE: 23-JAN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: COOPER, Iver P.
: REGISTRATION NUMBER: 28 005
: REFERENCE/DOCKET NUMBER: HANSSON-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202 628-5197
: TELEFAX: 202-737-3528
: INFORMATION FOR SEQ ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 857 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEetical: NO
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 45..593
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 45..593
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: 45..104
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 13..44
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 594..848
US-08-256-799-1
```

Query Match 9.4%, Score 58.6; DB 4; Length 857.
Best Local Similarity 67.8%; Prod. No. 00029;

```
Matches 82: Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 505 tcttttllactgaactatttaatttttttcaatttttcaacaaataaattatttttaagaa 564
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 719 tcttctctcttttaatttttaatttttttttttttttttttttttttttttttttttt 779
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 565 tgttaattgtattgaatttttttttttttttttttttttttttttttttttttttttt 624
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 779 AGCAACTGATTGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 838
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 625 a 625
Db 839 A 839

RESULT 4
US-08-462-437-1
: Sequence 1, Application: US/08462437
: Patent No. 6232094
: GENERAL INFORMATION:
: APPLICANT: HANSSON, Lennart
: APPLICANT: STROEMQVIST, Mats
: APPLICANT: BERGSTROEM, Sven
: APPLICANT: HERNELL, Olie
: APPLICANT: Toernell, Jan
: TITLE OF INVENTION: DNA ENCODING KAPPA-CASEIN, PROCESS
: TITLE OF INVENTION: FOR OBTAINING THE PROTEIN AND USE THEREOF
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWDY AND NEIMARK
: STREET: 419 Seventh Street, N.W., Suite 300
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/88/462,437
: FILING DATE: 05-JUN-1995
: PPI/P APPLICATION DATA:
: APPLICATION NUMBER: DF 88/92
: FILING DATE: 23-JAN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: COOPER, Iver P.
: REGISTRATION NUMBER: 28,005
: REFERENCE/DOCKET NUMBER: HANSSON-1A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: TELEX: 248633
: INFORMATION FOR SEQ ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 857 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEtical: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 45..593
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 45..593
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: 45..104
```


Lb 789 CATAAAGTAGATGATTTCCTGGGAACAATATATCTCTGACGACAAAGGATGT 848
Cy 498 cataatgtcttttactagactatttaatttttccaaattttccaaatatttttt 557
Lb 849 TCCTGTTTAAATTAATGACCTGTGATTTCCTGGTATGCTATTTAAATTTAACTA 908
Cy 558 aatggaatgttaattgaattgaattgaattgaattgaattgaattgaattgaatt 617
Lb 909 AAAAAAGTCTTCATCTTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 968
Cy 618 aaaaaaaa 625
Lb 969 AAAAAAAA 976

RESULT 7

US-09-247-373B 25
Sequence 25, Application US/09247373B
Patent No. 6168954

GENERAL INFORMATION:
APPLICANT: MCGONIGLE, BRIAN
APPLICANT: O'KEEFE, DANIEL
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
FILE REFERENCE: CL-1108-A
CURRENT APPLICATION NUMBER: US/09/247,373B
CURRENT FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 08/924,747
PRIOR FILING DATE: 1997-09-05
NUMBER OF SEQ ID NOS: 56
SOFTWARE: Microsoft Office 97
SEQ ID NO 25
LENGTH: 991
TYPE: DNA
ORGANISM: SOYBEAN

US-09-247-374B-25

Query Match 9.1%, Score 56.8, DB 4, Length 991,
Best Local Similarity 56.4%, Pred No. 0.00069,
Matches 106; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
Cy 438 catcaatgactgtatataaattgctcctaaaggagaaattctatgtatgagagaataat 497
Lb 789 cataaataatgaatgaatttcttggaacaattatctctgtatgagaaagattgt 848
Cy 498 cataaafcttttactagactatttaatttttccaaattttccaaatatttttt 557
Lb 849 tctgttttaatttaattgactgtgattgttggtggtattggtatttttaacta 908
Cy 558 aatggaatgttaattgaattgaattgaattgaattgaattgaattgaattgaatt 617
Lb 909 aaaaaatgtcaatttttaaaaaaataaaaaaataaaaaaataaaaaaataaaaaa 968
Cy 618 aaaaaaaa 625
Lb 969 aaaaaaaa 976

RESULT 8

US-09-296-715-25
Sequence 25, Application US/09296715
Patent No. 6171839

GENERAL INFORMATION:
APPLICANT: MCGONIGLE, BRIAN
APPLICANT: O'KEEFE, DANIEL
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
ENZYMES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: F.I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE

COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/296,715
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CL-1108
TELECOMMUNICATION INFORMATION:
TELEPHONE: 402-892-8112
TELEFAX: 402-773-0164
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 991 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
TISSUE TYPE: SOYBEAN
IMMEDIATE SOURCE:
CLONE: SSM.PK0067.G5
US-09-296-715-25

Query Match 9.1%, Score 56.8, DB 4, Length 991,
Best Local Similarity 56.4%, Pred No. 0.00069,
Matches 106; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
Cy 438 catcaatgactgtatataaattgctcctaaaggagaaattctatgtatgagagaataat 497
Lb 789 CATAAAGTAGATGATTTCCTGGGAACAATATATCTCTGACGACAAAGGATGT 848
Cy 498 cataaafcttttactagactatttaatttttccaaattttccaaatatttttt 557
Lb 849 TCCTGTTTAAATTAATGACCTGTGATTTCCTGGTATGCTATTTAAATTTAACTA 908
Cy 558 aatggaatgttaattgaattgaattgaattgaattgaattgaattgaattgaatt 617
Lb 909 AAAAAAGTCTTCATCTTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 968
Cy 618 aaaaaaaa 625
Lb 969 AAAAAAAA 976

RESULT 9

US-08-821-994-64
Sequence 64, Application US/08821994A
Patent No. 6228643
GENERAL INFORMATION:
APPLICANT: Greenland, Andrew J
APPLICANT: Thomas, Didier RP
APPLICANT: Jepson, Ian
TITLE OF INVENTION: Promoters
FILE REFERENCE: PPD 50108
CURRENT APPLICATION NUMBER: US/08/821,994A
CURRENT FILING DATE: 1997-03-23
EARLIER APPLICATION NUMBER: PCT/GB97/00729
EARLIER FILING DATE: 1997-03-18
EARLIER APPLICATION NUMBER: GB 9606062.9
EARLIER FILING DATE: 1996-03-22
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 64
LENGTH: 1474
TYPE: DNA
ORGANISM: Brassica napus
US-08-821-994-64

Query Match
Best Local Similarity 9.18; Score 56.6; DB 4; Length 1474;
Matches 83; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 499 ataatgttttttactgaactatttcaattttcaccacaaataagattatttta 558
DB 1298 ataatgttttttactgaactatttcaattttcaccacaaataagattatttta 1457
QY 559 atqqaatttttaattgaattgaataataaaaaaaataaaaaaaataaaaa 618
DB 1358 ataatgttttttactgaactatttcaattttcaccacaaataagattatttta 1417
QY 619 aaaaaa 625
DB 1418 aaaaaa 1424

RESULT 10
US-09-313-300-625
Sequence 6, Application US/09313300
Patent No. 622027
GENERAL INFORMATION:
APPLICANT: Kaser, Matthew, R.
APPLICANT: Lal, Preeti
APPLICANT: Yue, Henry
APPLICANT: Tang, Tom, Y.
APPLICANT: Baughn, Mariah, R.
APPLICANT: Azimzal, Yaida
TITLE OF INVENTION: MOLECULES EXPRESSED IN HIPPOCAMPOS
FILE REFERENCE: PR-0012 US
CURRENT APPLICATION NUMBER: US/09/313,300
CURRENT FILING DATE: 1999-05-17
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PPR Program
SEQ ID NO 6
LENGTH: 703
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY:
OTHER INFORMATION: 1941247
PUBLICATION INFORMATION:
US-09-313-300-6

Query Match
Best Local Similarity 9.08; Score 56.2; DB 4; Length 703;
Matches 82; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 501 atagtcttttttactgaactatttcaattttcaccacaaataagattatttta 560
DB 144 ATTGTACTGTTTCTCTAICTTCTATATGTTTAAAGTATATAATAAATATTATT 85
QY 561 ataatgttttttactgaactatttcaattttcaccacaaataagattatttta 620
DB 84 CTGTGTTTCACTTCATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 25
QY 621 aaaa 625
DB 24 AAAAA 20

RESULT 11
US-08-545-196B-10
Sequence 10, Application US/08545196B
Patent No. 608057

GENERAL INFORMATION:
APPLICANT: MELKI, JUDITH
APPLICANT: MURNICH, ARNOLD
TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-WIN/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,196B
FILING DATE: 19-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FARACI, C. J.
REGISTRATION NUMBER: 32,350
REFERENCE/DOCKET NUMBER: 2121-110P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1582 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-545-196B-10

Query Match
Best Local Similarity 9.08; Score 56.2; DB 3; Length 1582;
Matches 82; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 501 atagtcttttttactgaactatttcaattttcaccacaaataagattatttta 560
DB 1424 ATTGTACTGTTTCTCTAICTTCTATATGTTTAAAGTATATAATAAATATTATT 1483
QY 561 ataatgttttttactgaactatttcaattttcaccacaaataagattatttta 620
DB 1484 TTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1543
QY 621 aaaa 625
DB 1544 AAAAA 1548

RESULT 12
US-08-545-196B-12
Sequence 12, Application US/08545196B
Patent No. 608057
GENERAL INFORMATION:
APPLICANT: MELKI, JUDITH
APPLICANT: MURNICH, ARNOLD
TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747


```

QY 429 atctttcagcatcagctctgtatataaattgctcctaaagaaataatctatattgaca 489
DB 531 www. .... 590
QY 489 gaaataatacctataatcttttttctagtagctatttaatttttctagtagtagtag 548
DB 591 www. .... 650
QY 549 gattatttaataatgattatgattatgattatgattatgattatgattatgattatg 608
DB 651 www. .... 710
QY 609 aaaaaaaataaaaaaa 625
DB 711 www. .... 727

RESULT 9
AAF58254
ID AAF58254 standard; DNA; 936 BP.
XX
AC AAF58254;
XX
CT 24 APR-2001 (first entry)
XX
DE Oligonucleotide D1875.
XX
KW Electron-transfer group: ETM; mismatch; genotyping;
KW gene expression; ss.
XX
GS Synthetic.
XX
EN W0200107665-A2.
XX
PO 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WS-US20476.
XX
FR 26-JUL-1999; 990US-0145695.
PP 17-MAR-2000; 2000US-0190259.
XX
FA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Onek RM;
XX
DR WPT: 2001-159728/16
XX
PT Nucleic acids containing electron transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping; allowing repeat analyses on
PT a single surface.
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

```

```

Query Match 17.5%; Score 109.4; DB 22; Length 936;
Best Local Similarity 0.8%; Pred. No. 1,3e-10;
Matches 5; Conservative 93; Mismatches 219; indels 0; Gaps 0;

```

```

QY 9 atcaacagaaggtttaaagtaaaatctatttctattagaataatcagaaataatctactc 62
DB 111 www. .... 170
QY 63 ttaagatatacactagagagaggtatcgttttttttttttttttttttttttttttttt 128

```

```

DB 171 www. .... 230
QY 129 tcaacaagcagcagctctgtatataaattgctcctaaagaaataatctatattgaca 188
DB 231 www. .... 290
QY 189 aacataaactgattgacaaatcgcgcgttaacaagaattcgaqgtcuaatttaataaaca 248
DB 291 www. .... 350
QY 249 ttgtatttgacaaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 308
DB 351 www. .... 410
QY 409 aatagatgctctctctctctctctctctctctctctctctctctctctctctctctctct 358
DB 411 www. .... 470
QY 469 tgaagagctatttctcagctcattctctgagcattcgaatcgaatcgaatcgaatcgaatc 428
DB 471 www. .... 530
QY 429 atctttcagcatcagctctgtatataaattgctcctaaagaaataatctatattgaca 488
DB 531 www. .... 590
QY 489 gaaataatacctataatcttttttctagtagctatttaatttttctagtagtagtag 548
DB 591 www. .... 650
QY 549 gattatttaaaaggaattgattatgattatgattatgattatgattatgattatgattatg 608
DB 651 www. .... 710
QY 609 aaaaaaaataaaaaaa 625
DB 711 www. .... 727

```

```

RESULT 10
AAF58257
ID AAF58257 standard; DNA; 936 BP.
XX
AC AAF58257;
XX
CT 24 APR-2001 (first entry)
XX
DE Oligonucleotide D1954.
XX
KW Electron-transfer group: ETM; mismatch; genotyping;
KW gene expression; ss.
XX
GS Synthetic.
XX
EN W0200107665-A2.
XX
PO 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WS-US20476.
XX
FR 26-JUL-1999; 990US-0145695.
PP 17-MAR-2000; 2000US-0190259.
XX
FA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Onek RM;
XX
DR WPT: 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping; allowing repeat analyses on
PT a single surface.

```


^c and single-nucleotide polymorphisms, e.g. for genotyping. ^d monitoring gene expression.

Notes and references

Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match	17.5%;	Score 109.4;	DB 22;	Length 948;
Best Local Similarity	0.8%;	prod. No. 130-10;		
Matches	5;	Conservative 494;	Mismatches 219;	Indels 0;
				Gaps 0;

9 ut caaadadqdttaaqatqaadcatfcaatqaadqatcaaaadtttctc

69 114aqrataacagatgaagqgqgqqtattctttqatgqgtctctctctatgatt 128

[illegible][illegible][illegible][illegible]

411 www.wiley.com/go/4711

471

[illegible]

548

549 quttaitt taat iqaat q'ta at q'at ta q'aat q'aaat aaaaataaaataaaataaa 6108

```

06      651 ***** 719
07      609 aaaaaaaaaa 625

```

711 WWWWWW 727

RESULT 14
AAX19930

PP	04-FEB-1999.	
XX		
PF		
XX	24-JUL-1998;	98WO-AD000687.
XX		
PR	31-DEC-1997;	97AO-0001184.
PR	25-JUL-1997;	97AO-0008233.
XX		
PA	(OYME) UNIV MELBOURNE.	
XX		
PI	Bhalla P, Singh M, Swoboda L, Xu B;	
XX		
DR	WP1; 1999-142936/12.	
XX		
XX		
PT	New isolated nucleic acid tot., e.g. generating male sterile plants -	
PT	comprises a nucleotide sequence that allows expression of a	
PT	transposase gene in generative and sperm cells of a plant	
XX		
PS	Claim 8; Fig 10; 64pp; English.	

The present sequence represents the L3C1 promoter from Lily (Lilium longiflorum). The sequence specifically allows expression of the

CC genetic construct comprising a generative cell and sperm cell specific promoter operably linked to a transposase, which is capable of inducing transposition of a transposon, such that the transposon is integrated into the genome of the cell.

CC promoter, the transposase gene is expressed facilitating transposition of the element, (2) inducing male sterility in a plant by operably linking a CC cytotoxic nucleic acid to a promoter, which directs male-germ-specific

expression to produce a product from the cytoplasm and here and there inactivates or kills male gametes in the plant; and (5) a male sterile plant produced by the method. The male gamete promoter and/or gene is

CC the product of pollen grains is screened for a range of phenotypes of interest, and the transposon tagged plants are used to clone certain

reduced seed content, particularly where pollination stimulates fruit development, and where the lack of fertilisation results in seedless fruit. The nucleic acids enable specific genetic manipulation of the

CC male germ line by degrading male sterile plants, and facilitating male gamete specific transposon tagging.

XX

1. *Chlorophyll a* (Chl *a*)

```

XX OS Synthetic.
XX PN W0200107665-A2.
XX PD 01-FEB-2001.
XX PF 25-JUL-2000; 2000W0-US20476.
XX PR 25-JUL-1999; 99US-0145695.
XX DR 17-MAR-2000; 2000US-0190259.
XX PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX PT Umek RM.
XX DK WP1; 2001-159722/1b.
XX PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface.
XX PS Example 4; Page 120; 159pp; English.
XX CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETG) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX SQ Sequence 244 bp, 19 A, 9 C, 12 G, 10 T, 194 other;

Query March 11.69; Score 72.6; DS 22; Length 244;
Best Local Similarity 10.68; Pred. No. 0.00022;
Matches 21; Conservative 132; Mismatches 46; Indels 0; Gaps 0;

QY 427 ccactttcagratcagctctgatatagaattgcctcagggaggaattctagttq 486
Db 240 CCATTCGCGAGCTTCTCATTGATGCTCTTTTAAACAWWWWWWWWWWWWWWWWW 181
QY 487 cagaggaataatcatalagctctttttttatgaagctatttaatttttttga 546
Db 180 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 121
QY 547 aagattattttatggaatgttaattgaaataaaataaaataaaataaa 606
Db 120 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 61
QY 607 aaaaaaaataaaataaaataaaataaaataaaataaaataaaataaa 625
Db 60 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 42

```

Search completed: July 23, 2001, 01:43:45
 Job time: 2956 sec



GenCore version 4.5
Copyright (c) 1993 2000 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: July 22, 2001, 00:20:54 / Search time 1372.91 seconds
(without alignments)
4303 512 Million cell updates/sec

Title: US-09-463-480-3
Perfect score: 625
Sequence: 1 gccatccatcaacagaagg.....aaaaaaaaaaaaaaaaaaaaa 625

Scoring table: IDENTITY_NIC
Gapop 10.0, Gapext 1.0

Searched: 10229115 seqs, 4756426750 residues
Total number of hits satisfying chosen parameters: 23456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*
12: gb_est12.*
13: gb_est13.*
14: gb_est14.*
15: gb_est15.*
16: gb_est16.*
17: gb_est17.*
18: gb_est18.*
19: gb_est19.*
20: gb_est20.*
21: gb_est21.*
22: gb_est22.*
23: gb_est23.*
24: gb_est24.*
25: gb_est33.*
26: gb_est34.*
27: gb_est35.*
28: gb_est36.*
29: gb_est37.*
30: gb_est38.*
31: gb_est39.*
32: gb_est40.*
33: em_estba.*
34: em_estfun.*
35: em_esthum1.*
36: em_esthum2.*
37: em_esthum3.*
38: em_esthum4.*
39: em_esthum5.*
40: em_esthum6.*
41: em_esthum7.*
42: em_esthum8.*
43: em_esthum9.*

44: em_esthum10.*
45: em_esthum11.*
46: em_esthum12.*
47: em_esthum13.*
48: em_esthum14.*
49: em_esthum15.*
50: em_esthum16.*
51: em_esthum17.*
52: em_esthum18.*
53: em_esthum19.*
54: em_esthum20.*
55: em_esthum21.*
56: em_esthum22.*
57: em_esthum23.*
58: em_esthum24.*
59: em_esthum25.*
60: em_esthum26.*
61: em_esthum27.*
62: em_esthum28.*
63: em_estin1.*
64: em_estin2.*
65: em_estin3.*
66: em_estin4.*
67: em_estin5.*
68: em_estom1.*
69: em_estom2.*
70: em_estov1.*
71: em_estov2.*
72: em_estpl1.*
73: em_estpl2.*
74: em_estpl3.*
75: em_estpl4.*
76: em_estpl5.*
77: em_estpl6.*
78: em_estpl7.*
79: em_estpl8.*
80: em_estpl9.*
81: em_estpl10.*
82: em_estro1.*
83: em_estro2.*
84: em_estro3.*
85: em_estro4.*
86: em_estro5.*
87: em_estro6.*
88: em_estro7.*
89: em_estro8.*
90: em_estro9.*
91: em_estro10.*
92: em_estro11.*
93: em_estro12.*
94: em_estro13.*
95: em_estro14.*
96: em_estro15.*
97: em_estro16.*
98: em_estro17.*
99: em_estro18.*
100: em_estro19.*
101: em_estro20.*
102: gb_est25.*
103: gb_est26.*
104: gb_est27.*
105: gb_est28.*
106: gb_est29.*
107: gb_est30.*
108: gb_est31.*
109: gb_est32.*
110: gb_est41.*
111: gb_est42.*
112: gb_est43.*
113: gb_est44.*
114: gb_est45.*
115: gb_est46.*
116: gb_est47.*

117: qb_est148.*
 118: qb_est149.*
 119: qb_est150.*
 120: qb_est151.*
 121: qb_est152.*
 122: qb_est153.*
 123: qb_est154.*
 124: qb_est155.*
 125: qb_est156.*
 126: qb_est157.*
 127: qb_est158.*
 128: qb_est159.*
 129: qb_est160.*
 130: qb_est161.*
 131: qb_est162.*
 132: qb_est163.*
 133: qb_est164.*
 134: qb_est165.*
 135: qb_est166.*
 136: qb_est167.*
 137: qb_est168.*
 138: qb_est169.*
 139: qb_est170.*
 140: qb_est171.*
 141: qb_est172.*
 142: qb_est173.*
 143: qb_est174.*
 144: qb_est175.*
 145: qb_est176.*
 146: qb_est177.*
 147: qb_est178.*
 148: qb_est179.*
 149: qb_est180.*
 150: qb_est181.*
 151: qb_est182.*
 152: qb_est183.*
 153: qb_est184.*
 154: qb_est185.*
 155: qb_est186.*
 156: qb_est187.*
 157: qb_est188.*
 158: qb_est189.*
 159: qb_est190.*
 160: qb_est191.*
 161: qb_est192.*
 162: qb_est193.*
 163: qb_est194.*
 164: qb_est195.*
 165: qb_est196.*
 166: qb_est197.*
 167: qb_est198.*
 168: qb_est199.*
 169: qb_est200.*
 170: qb_est201.*
 171: qb_est202.*
 172: qb_est203.*
 173: qb_est204.*
 174: qb_est205.*
 175: qb_est206.*
 176: qb_est207.*
 177: qb_est208.*
 178: qb_est209.*
 179: qb_est210.*
 180: qb_est211.*
 181: qb_est212.*
 182: qb_est213.*
 183: qb_est214.*
 184: qb_est215.*
 185: qb_est216.*
 186: qb_est217.*
 187: qb_est218.*
 188: qb_est219.*
 189: qb_est220.*

190: qb_est221.*
 191: qb_est222.*
 192: qb_est223.*
 193: qb_est224.*
 194: qb_est225.*
 195: qb_est226.*
 196: qb_est227.*
 197: qb_est228.*
 198: qb_est229.*
 199: qb_est230.*
 200: qb_est231.*
 201: qb_est232.*
 202: qb_est233.*
 203: qb_est234.*
 204: qb_est235.*
 205: qb_est236.*
 206: qb_est237.*
 207: qb_est238.*
 208: qb_est239.*
 209: qb_est240.*
 210: qb_est241.*
 211: qb_est242.*
 212: qb_est243.*
 213: qb_est244.*
 214: qb_est245.*
 215: qb_est246.*
 216: qb_est247.*
 217: qb_est248.*
 218: qb_est249.*
 219: qb_est250.*
 220: qb_est251.*
 221: qb_est252.*
 222: qb_est253.*
 223: qb_est254.*
 224: qb_est255.*
 225: qb_est256.*
 226: qb_est257.*
 227: qb_est258.*
 228: qb_est259.*
 229: qb_est260.*
 230: qb_est261.*
 231: qb_est262.*
 232: qb_est263.*
 233: qb_est264.*
 234: qb_est265.*
 235: qb_est266.*
 236: qb_est267.*
 237: qb_est268.*
 238: qb_est269.*
 239: qb_est270.*
 240: qb_est271.*
 241: qb_est272.*
 242: qb_est273.*
 243: qb_est274.*
 244: qb_est275.*
 245: qb_est276.*
 246: qb_est277.*
 247: qb_est278.*
 248: qb_est279.*
 249: qb_est280.*
 250: qb_est281.*
 251: qb_est282.*
 252: qb_est283.*
 253: qb_est284.*
 254: qb_est285.*
 255: qb_est286.*
 256: qb_est287.*
 257: qb_est288.*
 258: qb_est289.*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

QY 512 1actqagctatfataattttcaatttccaccataagaattatttttaagguaatqttaat 571

located
edge
site of
then
rest,
bottom of

545

278

— — —

112

system;

and cDNA



Matches	32; Conservative	35; Mismatches	52; Indels	47; Gaps	5; 5;

```

QY 10 CVLFVWVKAAALADUKVNP-----TQPMVTQTITITIGGKQEEF 50
   : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 QAVQIVNNAVLDPNCFEYENMCEQLKELVGLSPTEVFLFAAPASTITAIKFAHE 374
   : : : : : : : : : : : : : : : : : : : : : : : :
QY 51 VNLINLCAOSNVKVSCD-----GLHTEPDPHITRPLSDGTNNKLVNG---APL 100
   : : : : : : : : : : : : : : : : : : : : : : : :
Db 375 AINITWSECSVESFVGKDGKQEGEDYSANEIE-HVLPQVIEGTRKVLGNGLGMVIL 433
   : : : : : : : : : : : : : : : : : : : : : : : :
QY 101 SHATLVAF-----KYANDVPESFSI 120
   : : : : : : : : : : : : : : : : : : : : : : : :
Db 434 INGLISLQNNMTWNCKGFGTAPSTPI 460
   : : : : : : : : : : : : : : : : : : : : : : : :

```

```

1 RESULT
2
3 US-08-872-855-11
4
5 : Sequence 11, Application US/08872855
6
7 : Patent No. 6121045
8
9 : GENERAL INFORMATION:
10
11 : APPLICANT: McCarthy, Sean
12
13 : APPLICANT: Garfield, David
14
15 : TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
16
17 : TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
18
19 : NUMBER OF SEQUENCES: 23
20
21 : CORRESPONDENCE ADDRESS:
22
23 : ADDRESSEE: FOLEY, HOAG & ELIOT LLP
24
25 : STREET: One Post Office Square
26
27 : CITY: Boston
28
29 : STATE: MA
30
31 : COUNTRY: USA
32
33 : ZIP: 02109-2170
34
35 : COMPUTER READABLE FORM:
36
37 : MEDIUM TYPE: Floppy disk
38
39 : COMPUTER: IBM PC compatible
40
41 : OPERATING SYSTEM: PC-DOS/MS-DOS
42
43 : SOFTWARE: Patented to Applicant #1 6, Applicant #1 30
44
45 : CURRENT APPLICATION DATA:
46
47 : APPLICATION NUMBER: US/08/872,855
48
49 : FILING DATE: 11-JUN-1997
50
51 : CLASSIFICATION: 514
52
53 : ATTORNEY/AGENT INFORMATION:
54
55 : NAME: Arnold, Beth E.
56
57 : REGISTRATION NUMBER: 35,430
58
59 : REFERENCE/STAKEI NUMBER: MAA-003-02
60
61 : TELECOMMUNICATION INFORMATION:
62
63 : TELEPHONE: 617-832-1060

```

Query Match 10.1%; Score 69; DB 4; Length 830;
Best Local Similarity 24.6%; Pred. No. 12;

QY	1	MPNAVFVAVVWVFWVKAALADK	-----TGNPTDPMVTOTITGLTIGG	45
DB		: : : : : : : : : :		
DB	603	MPVLVIAAANVFFPMKKRKAKGKDA	AKPKNQNAVATMHHNGSGVGLASASLG	662
QY	46	KQFFVNIININYYACSNVVFQAN	HTTFTTHHLEPTSPQNNIVVNEAFT	705
DB	663	K-----	: : : : : : : : : :	
		-----TGSNSLTDFG	-----GNPNIKNTWKSVNNICASAAAAA	703
QY	106	VA	107	
DB	704	AA	705	

RESULT 3
 US-08-264-534-6
 ? Sequence 6, Application US-08264534
 ? Patent No. 5648464
 ? GENERAL INFORMATION:
 ? APPLICANT: Artavanis-Tsakonas, Spyridon et al.
 ? TITLE OF INVENTION: In Topothymic Proteins, And Methods Based Thereon
 ? NUMBER OF SEQUENCES: 44
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Pennie & Edmonds
 ? STREET: 1155 Avenue of the Americas
 ? CITY: New York
 ? STATE: New York
 ? COUNTRY: U.S.A.
 ? ZIP: 10036
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: Patent In Release #1.0, Version #1.25
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US-08/264,534
 ? FILING DATE:
 ? CLASSIFICATION: 435
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 07/695,189
 ? FILING DATE: 03-MAY-1991
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Mistock, S. Leslie
 ? REGISTRATION NUMBER: 18,472
 ? REFERENCE/DOC# NUMBER: 7336-904
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: 212 390-9090
 ? TELEFAX: 212 390-8864/9741
 ? TELEX: 66141 PENNIE
 ? INFORMATION FOR SEQ ID NO: 6:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 833 amino acids
 ? TYPE: amino acid
 ? TOPOLOGY: unknown
 ? MOLECULE TYPE: protein
 ? US-08-264-534-6

Query Match 10.1%; Score 69; DB 1; Length 833;
Best Local Similarity 24.6%; Pred. No. 12;
Matches 30; Conservative 15; Mismatches 43; Indels 34; Gaps

QY	1	MRAVAVFACVLFCVMVHKAALADK-----TCNPTDFPMVTQTITGLTIGG	45
Db	605	MPILVAIAACVFVKKKPKAKFKTAAPKQNEGNATVMHHNSGVVALASAGI	664
QY	46	KQEEFNLLNNLYCAQSNVKSVDGLHTTEPDPHIIRPLSDGTNCLVNNGAIPISHATL	105
Db	665	K-----TGSNSGLTFDG-----GNPIIKNTWKSVINNICASAAAAAAA	705
QY	106	VA	107
Db	706	AA	707

RESULT 4
 US-08-083-590A-2
 ; Sequence 2, Affiliation US-08-083-590A
 ; Patent No. 5786158
 ; GENERAL INFORMATION:
 ; APPLICANT: Artavanis-Tsakonas, S. et al.
 ; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
 ; TITLE OF INVENTION: And Compositions Based On No. 5786158ch. proteins
 ; TITLE OF INVENTION: Nucleic Acids

: NUMBER OF SEQUENCES: 21
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennie & Edmonds
 : STREET: 1155 Avenue of the Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036
 : COMPUTER FEASIBLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/463,500A
 : FILING DATE: 25-JUN-1993
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Mistrock, S. Leslie
 : REGISTRATION NUMBER: 18,872
 : REFERENCE/DOCKET NUMBER: 7326-015
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 212 790-9090
 : TELEFAX: 212 8698864/9741
 : TELEX: 66141 PENNIE
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 833 amino acids
 : TYPE: amino acid
 : TOPOLOGY: unknown
 : MOLECULE TYPE: protein
 : US-08-083-590A-2

Query Match 10.1%, Score 69, DB 1, Length 833;
 Best Local Similarity 24.6%, Pred. No. 12;
 Matches 30; Conservative 15; Mismatches 43; Indels 34; Gaps 3;
 QY 1 MRVAVFFACVLCMVHKAALADK-----TCNPTDFMVTOTITGLTIGG 45
 Db 605 MPLVAVIAACVVFCKPKPKPAQKTAFAEKQNE_NAVATMHHSNGVGVVALASASL22 664
 QY 46 KQFEFVNLIINLYCAOSNVKVSQGLITTEPDPHILRLSDGTRNCLVHNGAPISHATL 105
 Db 665 K-----TGSNSGLTFDG-----GNPNLIKNWDSVNNICASAAAAAAA 705
 QY 106 VA 107
 Db 706 AA 707

RESULT 5
 US-08-465-500-6
 : Sequence 6, Application US/08/465,500
 : Patent No. 5789195
 : GENERAL INFORMATION:
 : APPLICANT: Artavanis-Tsakonas, Spyridon
 : APPLICANT: Muskavitch, Marc A. T.
 : APPLICANT: Fehon, Richard G.
 : APPLICANT: Rebay, Ilaria
 : APPLICANT: Blauvellet, Cristine M.
 : APPLICANT: Shephard, Scott B.
 : TITLE OF INVENTION: HUMAN NUCLEO AND DELTA, BINDING DOMAINS
 : TITLE OF INVENTION: IN TOPOTHYLMIC PROTEINS, AND METHODS BASED THEREON
 : NUMBER OF SEQUENCES: 34
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: PENNIE & EDMONDS
 : STREET: 1155 Avenue of the Americas
 : CITY: New York
 : STATE: NY
 : COUNTRY: USA
 : ZIP: 10036-2711
 : COMPUTER FEASIBLE FORM:

: MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/465,500
 : FILING DATE: 05-JUN-1995
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Mistrock, S. Leslie
 : REGISTRATION NUMBER: 18,872
 : REFERENCE/DOCKET NUMBER: 7326-034
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (212) 790-9090
 : TELEFAX: (212) 869-8864/9741
 : INFORMATION FOR SEQ ID NO: 6:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 833 amino acids
 : TYPE: amino acid
 : TOPOLOGY: unknown
 : MOLECULE TYPE: protein
 : US-08-465-500-6

Query Match 10.1%, Score 69; DB 1; Length 833;
 Best Local Similarity 24.6%, Pred. No. 12;
 Matches 30; Conservative 15; Mismatches 43; Indels 34; Gaps 3;
 QY 1 MRVAVFFACVLCMVHKAALADK-----TCNPTDFMVTOTITGLTIGG 45
 Db 605 MPLVAVIAACVVFCKPKPKPAQKTAFAEKQNE_NAVATMHHSNGVGVVALASASL22 664
 QY 46 KQFEFVNLIINLYCAOSNVKVSQGLITTEPDPHILRLSDGTRNCLVHNGAPISHATL 105
 Db 665 K-----TGSNSGLTFDG-----GNPNLIKNWDSVNNICASAAAAAAA 705
 QY 106 VA 107
 Db 706 AA 707

RESULT 6
 US-08-466-126-6
 : Sequence 6, Application US/08/466,126
 : Patent No. 5849869
 : GENERAL INFORMATION:
 : APPLICANT: Artavanis-Tsakonas, Spyridon et al.
 : TITLE OF INVENTION: Human No. 5849869ch And Delta, Binding Domains
 : TITLE OF INVENTION: In Topothylmic Proteins, And Methods Based Thereon
 : NUMBER OF SEQUENCES: 30
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennie & Edmonds
 : STREET: 1155 Avenue of the Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036
 : COMPUTER FEASIBLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/466,126
 : FILING DATE:
 : CLASSIFICATION: 530
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/791,923
 : FILING DATE: 14-NOV-1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Mistrock, S. Leslie
 : REGISTRATION NUMBER: 18,872
 : REFERENCE/DOCKET NUMBER: 7326-007

US-08-346-128-6

DB 605 MP1VAV1AACVVE7MKFKFKFAQCFK7FAFAFKQNFQNAVATMIHNGSGVGVALASASLGG 664

ADDRESSEE: BROWDY AND NEWMARK
STREET: 419 Seventh Street, Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/092.817
FILING DATE: 19-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 1L 100696
FILING DATE: 19-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 1L 102915
FILING DATE: 23-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: TOWNSEND, G. Kevin
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: RUBINSTEIN-5A
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633

Query Match 9.9%; Score 68; DB 1; Length 860;
Best Local Similarity 24.7%; Pred. No. 16;
Matches 38; Conservative 13; Mismatches 57; Indels 46; Gaps 8;
QY 8 FACVLCMVHKAALD-----DK-----TCNPTDPMVTQITGLFISGRQEF--- 49
DB 202 FHCLSGECIHSSWPCUGGPGDKKSTFTNCVAVATCPHDEF---QCSGNCIBHSRQCTURF 258
QY 50 -----EVNLIINLYCAQSNVKSVCGLHTTEPIDPHII-----RPLSD-----G 88
DB 259 YCKKMSDEVCVNVTLCEGPN-KERC---HSGECITLTKVNMARCKRQWSDEPIKECG 314
QY 89 TNNCLVNNCAPISHATLVAFKYAVDVPFSFIS 122
DB 315 TNECLNNGGCSHVCLNKGVECLCPDGQFOLVA 348

Query Match 9.9%; Score 68; DB 1; Length 860;
Best Local Similarity 24.7%; Pred. No. 16;
Matches 38; Conservative 13; Mismatches 57; Indels 46; Gaps 8;

QY 8 FACVLCMVHKAALD-----DK-----TCNPTDPMVTQITGLFISGRQEF--- 49
DB 202 FHCLSGECIHSSWPCUGGPGDKKSTFTNCVAVATCPHDEF---QCSGNCIBHSRQCTURF 258
QY 50 -----EVNLIINLYCAQSNVKSVCGLHTTEPIDPHII-----RPLSD-----G 88
DB 259 YCKKMSDEVCVNVTLCEGPN-KERC---HSGECITLTKVNMARCKRQWSDEPIKECG 314
QY 89 TNNCLVNNCAPISHATLVAFKYAVDVPFSFIS 122
DB 315 TNECLNNGGCSHVCLNKGVECLCPDGQFOLVA 348

RESULT 12
US-08-470-058-2
Sequence 2, Application US/08470058
Patent No. 5817789
GENERAL INFORMATION:
APPLICANT: Heartline, Michael W.
APPLICANT: Lemontt, Jeffrey P.
TITLE OF INVENTION: Chimeric Proteins For Use in Transport
TITLE OF INVENTION: of a Selected Substance Into Cells
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brock, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470.058
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: TKT93-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1074 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-470-058-2

Query Match 9.9%; Score 68; DB 2; Length 1074;
Best Local Similarity 24.7%; Pred. No. 22;
Matches 38; Conservative 13; Mismatches 57; Indels 46; Gaps 8;
QY 8 FACVLCMVHKAALD-----DK-----TCNPTDPMVTQITGLFISGRQEF--- 49
DB 202 FHCLSGECIHSSWPCUGGPGDKKSTFTNCVAVATCPHDEF---QCSGNCIBHSRQCTURF 258
QY 50 -----EVNLIINLYCAQSNVKSVCGLHTTEPIDPHII-----RPLSD-----G 88
DB 259 YCKKMSDEVCVNVTLCEGPN-KERC---HSGECITLTKVNMARCKRQWSDEPIKECG 314
QY 89 TNNCLVNNCAPISHATLVAFKYAVDVPFSFIS 122
DB 315 TNECLNNGGCSHVCLNKGVECLCPDGQFOLVA 348

RESULT 13
US-09-037-188-2
Sequence 2, Application US/09037188
Patent No. 6027921
GENERAL INFORMATION:
APPLICANT: Heartline, Michael W.
APPLICANT: Lemontt, Jeffrey P.
APPLICANT: Concino, Michael P.
TITLE OF INVENTION: CHIMERIC PROTEINS FOR USE
TITLE OF INVENTION: IN TRANSPORT OF A SELECTED SUBSTANCE INTO CELLS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037.188
FILING DATE: 02-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Ph.D., J.D., Janis K.

QY . 50 -----EVLINLYTASNVKVSIGLHTEHEDHII-----RPLSH-----G 88
DB 250 YKRNCEVGVNVTLEGERDEK-----ESGELLEEVSMARQETRAWLEFEDM 214
QY 89 INNVLNNGAPISHALIVAFKYAMVPOSESTIS 122
DB 315 TNECLNNNGSIVNDEKIGYEELCPDQFQIVA 448

Search completed: July 18, 2001, 17:28:17
Job time: 49 sec

GenCode version 4.5
Copyright (c) 1999 - 2000 Compugen Ltd

OM protein protein search, using sw model

Run on: July 18, 2001, 17:27:28 Search time 21.18 seconds
(without alignments)
361 260 million cell updates/sec

Title: US-09-463-480-4

Perfect score: 685

Sequence: 1 MPAAVAVFAACVLCVQVHKA

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 412676 seqs, 6062988 residues

Total number of hits satisfying chosen parameters: 412576

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_0601**

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT**

2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT**

3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT**

4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT**

5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT**

6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT**

7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT**

8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT**

9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT**

10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT**

11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT**

12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT**

13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT**

14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT**

15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT**

16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT**

17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT**

18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT**

19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT**

20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT**

21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT**

22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	685	100.0	128	AA04129	117 L1 protein
2	116.5	17.0	1048	AA04129	Arabidopsis thaliana
3	116.5	17.0	1129	AA04129	Arabidopsis thaliana
4	116.5	17.0	1193	AA04129	Arabidopsis thaliana
5	114.5	16.7	135	AA01099	Arabidopsis thaliana
6	114	16.6	124	AA04192	Arabidopsis thaliana
7	113.5	16.6	127	AA011090	Arabidopsis thaliana
8	95.5	13.9	70	AA027387	Arabidopsis thaliana
9	95.5	13.9	75	AA011091	Arabidopsis thaliana
10	95.5	13.9	75	AA027387	Arabidopsis thaliana
11	78.5	11.5	419	AA093100	WO 99/05785, 6-117

12	76.5	11.2	264	AA023735	Arabidopsis thaliana
13	76.5	11.2	261	AA023734	Arabidopsis thaliana
14	76.5	11.2	297	AA023733	Arabidopsis thaliana
15	76.5	11.2	317	AA049594	Arabidopsis thaliana
16	76.5	11.2	317	AA049625	Arabidopsis thaliana
17	76.5	11.2	354	AA049593	Arabidopsis thaliana
18	76.5	11.2	354	AA049624	Arabidopsis thaliana
19	76.5	11.2	354	AA049592	Arabidopsis thaliana
20	76.5	11.2	492	AA019623	Arabidopsis thaliana
21	76	11.1	422	AA033559	WO 99/07855 SeqID
22	73.5	10.7	316	AA015590	Arabidopsis thaliana
23	73.5	10.7	536	AA015589	Arabidopsis thaliana
24	73	10.7	2383	AA015945	E. coli proliferat
25	71	10.4	145	AA094865	ber p II antigen f
26	69.5	10.1	800	AA077114	human low density
27	69	10.1	130	AA033923	Arabidopsis thaliana
28	69	10.1	833	AA028960	delta D11. Homo s
29	68	9.9	750	AA047868	human f11 receptor
30	68	9.9	939	AA049601	human low density
31	68	9.9	860	AA047157	Sequence of human
32	68	9.9	860	AA047860	human LDL receptor
33	68	9.9	924	AA078234	chicken p55/human
34	68	9.9	1074	AA007621	LDLR/TF chimeric p
35	68	9.9	1410	AA007622	LDLR/TF chimeric p
36	67	9.8	408	AA019760	pyruvate:ferredoxi
37	67	9.8	519	AA018308	Plasmodium falcipa
38	67	9.8	834	AA079031	Drosophila delta p
39	67	9.8	1395	AA013563	Drosophila Pabo l
40	67	9.8	1395	AA008401	Drosophila sp. ROB
41	66.5	9.7	322	AA022599	l1b1 receptor fragm
42	66.5	9.7	419	AA041074	Amino acid sequenc
43	66.5	9.7	541	AA022876	YFP-NLS-CP3-multip
44	66.5	9.7	541	AA079654	Caspase-3 biosenso
45	66.5	9.7	6797	AA041598	Pinat c10n biosynth

ALIGNMENTS

RESULT	1
AA04129	
ID	AA04129 standard: Protein; 128 AA.
XX	AC AA04129;
XX	11-MIN-1999 (first entry)
XX	July 1991 protein sequence.
XX	Lily, Lilium longiflorum, f10L, f10A, g003, pollen, male; germ line;
KW	protein, sterile, transposase; sperm cell; pollination, seedless fruit;
RW	genetic manipulation, gamete, histone.
XX	Lilium longiflorum.
XX	WO9905281-A1.
PD	04-FEB-1999.
XX	24-JUL-1998; 98WO-A000587.
XX	31-DEC-1997; 97AO-0001184.
PF	26-MAR-1997; 97AB-B008233.
XX	(OYME) UNIV MELBOURNE.
XX	Bhalla P, Singh M, Swoboda L, Xu H;
XX	WP; 1999-142936/12.
XX	N-PSDB; AAX19927.
XX	New isolated nucleic acid for, e.g. generating male sterile plants -
XX	comprises a nucleotide sequence that allows expression of a

PT transposase gene in generative and sperm cells of a plant.
 XX Claim 4: Fig 1: 64pp; English.
 XX
 CC The present sequence represents the LGC1 protein from lily (Lilium
 CC longiflorum). The sequence specifically allows expression of the
 CC gene in generative and sperm cells of a plant. Also described are: (1) a
 CC genetic construct comprising a generative cell and sperm cell specific
 CC promoter operably linked to a transposase, which is capable of inducing
 CC transposition of a transposable element such that upon expression of the
 CC promoter, the transposase gene is expressed facilitating transposition of the
 CC element; (2) inducing male sterility in a plant by operably linking a
 CC cytotoxic nucleic acid to a promoter, which directs male gamete specific
 CC expression to produce a product from the cytotoxic nucleic acid that
 CC inactivates or kills male gametes in the plant; and (3) a male sterile
 CC plant produced by the method. The male gamete promoter and/or gene is
 CC useful for facilitating male gamete specific transposon tagging, where
 CC the product of pollen grains is screened for a range of phenotypes of
 CC interest, and the transposon tagged plants are used to clone certain
 CC genes. The method is useful for producing seedless fruit or fruit with
 CC reduced seed content, particularly where pollination stimulates fruit
 CC development, and where the lack of fertilisation results in seedless
 CC fruit. The nucleic acids enable specific genetic manipulation of the
 CC male germ line by generating male sterile plants, and facilitating male
 CC gamete specific transposon tagging
 XX
 SQ Sequence 128 AA;

Query Match 100.0%; Score 685; DB 20; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1,25 70;
 Matches 128, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
 QY 1 MAAVAVFAACVLFMVHKAALADKDFNPFHLMVPELLESLDSEKSEHVNLEMLVTA 63
 DB 1 mvaavafacvlfmvhkaaladkdnfptfntvtrtqtatfagggfevefalmvta 60
 QY 61 QSNVKSVDGLHTIERIDPHILRLSECTNNLNNNGAPISHAILEVAFKVAWVPPESL 120
 DB 61 qsnvksvdglhtieridphilrldsectnnl nnnngapisshailvafkvaawvppesf 120
 QY 121 ISSDINGS 128
 DB 121 issdins 128

RESULT 2
 AAC42154 ID AAC42154 standard; Protein: 1048 AA.
 XX AC AAC42154;
 XX DT 18-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 52635.
 XX KW Protein identification, signal transduction pathway, metabolic pathway,
 KW hybridisation assay, genetic mapping, gene expression control, promoter,
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PO 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0401440.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0121180.
 PR 23-MAR-1999; 99US-0124548.
 PR 03-MAR-1999; 99US-0125789.
 PR 25-MAR-1999; 99US-0126264.

PF 29 MAR 1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 23-APR-1999; 99US-0131444.
 PR 30-APR-1999; 99US-0132048.
 PR 22 APR 1999; 99US-0132407.
 PR 04 MAY 1999; 99US-0132484.
 PR 05 MAY 1999; 99US-0132485.
 PR 06 MAY 1999; 99US-0132486.
 PR 06 MAY 1999; 99US-0132487.
 PR 07 MAY 1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0132856.
 PR 14-MAY-1999; 99US-0134218.
 PR 14 MAY 1999; 99US-0134219.
 PR 14 MAY 1999; 99US-0134221.
 PR 14 MAY 1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 31-JUN-1999; 99US-0137222.
 PR 33 JUN 1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141247.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 15-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325
 PR 19-JUL-1999; 99US-0144331
 PR 19-JUL-1999; 99US-0144332
 PR 19-JUL-1999; 99US-0144333
 PR 19-JUL-1999; 99US-0144334
 PR 19-JUL-1999; 99US-0144335
 PR 20-JUL-1999; 99US-0144352
 PR 20-JUL-1999; 99US-0144632
 PR 20-JUL-1999; 99US-0144884
 PR 21-JUL-1999; 99US-0144814
 PR 21-JUL-1999; 99US-0145086
 PR 21-JUL-1999; 99US-0145088
 PR 22-JUL-1999; 99US-0145088
 PR 22-JUL-1999; 99US-0145087
 PR 22-JUL-1999; 99US-0145089
 PR 22-JUL-1999; 99US-0145192
 PR 23-JUL-1999; 99US-0145145
 PR 23-JUL-1999; 99US-0145218
 PR 23-JUL-1999; 99US-0145224
 PR 23-JUL-1999; 99US-0145274
 PR 23-JUL-1999; 99US-0145276
 PR 27-JUL-1999; 99US-0145913
 PR 27-JUL-1999; 99US-0145918
 PR 27-JUL-1999; 99US-0145919
 PR 28-JUL-1999; 99US-0145951
 PR 02-AUG-1999; 99US-0146386
 PR 02-AUG-1999; 99US-0146388
 PR 02-AUG-1999; 99US-0146389
 PR 03-AUG-1999; 99US-0147038
 PR 04-AUG-1999; 99US-0147204
 PR 04-AUG-1999; 99US-0147302
 PR 05-AUG-1999; 99US-0147192
 PR 05-AUG-1999; 99US-0147260
 PR 06-AUG-1999; 99US-0147403
 PR 06-AUG-1999; 99US-0147416
 PR 09-AUG-1999; 99US-0147493
 PR 09-AUG-1999; 99US-0147935
 PR 10-AUG-1999; 99US-0148171
 PR 11-AUG-1999; 99US-0148319
 PR 12-AUG-1999; 99US-0148431
 PR 13-AUG-1999; 99US-0148565
 PR 13-AUG-1999; 99US-0148684
 PR 16-AUG-1999; 99US-0149368
 PR 17-AUG-1999; 99US-0149175
 PR 18-AUG-1999; 99US-0149426
 PR 20-AUG-1999; 99US-0149722
 PR 20-AUG-1999; 99US-0149723
 PR 20-AUG-1999; 99US-0149929
 PR 23-AUG-1999; 99US-0149902
 PR 23-AUG-1999; 99US-0149930
 PR 25-AUG-1999; 99US-0150566
 PR 26-AUG-1999; 99US-0150884
 PR 27-AUG-1999; 99US-0151065
 PR 27-AUG-1999; 99US-0151066
 PR 27-AUG-1999; 99US-0151080
 PR 30-AUG-1999; 99US-0151303
 PR 31-AUG-1999; 99US-0151438
 PR 01-SEP-1999; 99US-0151930
 PR 07-SEP-1999; 99US-0152664
 PR 10-SEP-1999; 99US-0153070
 PR 13-SEP-1999; 99US-0153758
 PR 15-SEP-1999; 99US-0154018
 PR 16-SEP-1999; 99US-0154039
 PR 20-SEP-1999; 99US-0154779
 PR 22-SEP-1999; 99US-0155139
 PR 23-SEP-1999; 99US-0155486
 PR 24-SEP-1999; 99US-0155654
 PR 28-SEP-1999; 99US-0156458
 PR 29-SEP-1999; 99US-0156596
 PR 04-OCT-1999; 99US-0157117
 PR 05-OCT-1999; 99US-0157753
 PR 06-OCT-1999; 99US-0157865
 PR 07-OCT-1999; 99US-0158029
 PR 08-OCT-1999; 99US-0158242

PR 13-OCT-1999; 99US-0158669
 PR 13-OCT-1999; 99US-0159293
 PR 13-OCT-1999; 99US-0159294
 PR 13-OCT-1999; 99US-0159295
 PR 14-OCT-1999; 99US-0159329
 PR 14-OCT-1999; 99US-0159330
 PR 14-OCT-1999; 99US-0159331
 PR 14-OCT-1999; 99US-0159637
 PR 14-OCT-1999; 99US-0159638
 PR 14-OCT-1999; 99US-0159584
 PR 21-OCT-1999; 99US-0160741
 PR 21-OCT-1999; 99US-0160767
 PR 21-OCT-1999; 99US-0160768
 PR 21-OCT-1999; 99US-0160814
 PR 21-OCT-1999; 99US-0160815
 PR 25-OCT-1999; 99US-0160980
 PR 25-OCT-1999; 99US-0160981
 PR 25-OCT-1999; 99US-0160989
 PR 25-OCT-1999; 99US-0161404
 PR 25-OCT-1999; 99US-0161405
 PR 26-OCT-1999; 99US-0161406
 PR 26-OCT-1999; 99US-0161359
 PR 26-OCT-1999; 99US-0161360
 PR 26-OCT-1999; 99US-0161361
 PR 28-OCT-1999; 99US-0161920
 PR 28-OCT-1999; 99US-0161992
 PR 28-OCT-1999; 99US-0161993
 PR 29-OCT-1999; 99US-0162142

Query Match

17.0%, Score 116.5, DB 21, Length 1048,

Best local similarity 27.8%,

Prod. No. 0.00917,

Matches 35; Conservative 17, Mismatches 51;

Indels 23; Gaps 5;

QY 3 AVAFVACV-----LFCMVHKAALADUKICNPIDFMVIGIITGLTIGCK 46

DB 912 asvqllcvymstpkllllwssialillqsgalceaksscpa...vqlqykwiekk 968

QY 17 QEEVRLRLNLVCASENVKVSIGLHLEPILFHLFELSGINRCLVNRGAPISHATL 105

DB 969 --frvcmkpcmqclirikeqgtf-pqstvdqpftrvsssqmcrvndgipispqyt 1025

QY 106 VAFKYA III

LB 1026 lsfny 1031

RESULT 3

AAG12153

ID AAG12153 standard; Protein, 1123 AA.

XX

AC AAG12153;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 52534.

XX

KW protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX

PC 06-SEP-2000.

XX

PF 25-FEB-2000; 99US-0161439.

XX

PP 25-FEB-1999; 99US-0121825.

PP 05-MAP-1999; 99US-0174180.

PP 09-MAP-1999; 99US-0123548.

PP 23-MAP-1999; 99US-0125788.

PR 25-MAR-1999; 990S-0126254.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130444.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132437.
PR 04-MAY-1999; 990S-0132494.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132436.
PR 06-MAY-1999; 990S-0132447.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 07-JUN-1999; 990S-0137532.
PR 08-JUN-1999; 990S-0137724.
PR 10-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139442.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139753.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139849.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140923.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144341.
PR 19-JUL-1999; 990S-0144342.
PR 19-JUL-1999; 990S-0144343.
PR 19-JUL-1999; 990S-0144344.
PR 19-JUL-1999; 990S-0144345.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 23-JUL-1999; 990S-0145142.
PR 23-JUL-1999; 990S-0145145.
PR 24-JUL-1999; 990S-0145218.
PR 24-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148419.
PR 12-AUG-1999; 990S-0148441.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149910.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151348.
PR 01-SEP-1999; 990S-0151940.
PR 07-SEP-1999; 990S-0152463.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.

```

PR 08-OCT-1999; 990S-0159232
PR 12-OCT-1999; 990S-0159264
PR 13-OCT-1999; 990S-0159293
PR 13-OCT-1999; 990S-0159264
PR 13-OCT-1999; 990S-0159295
PR 14-OCT-1999; 990S-0159329
PR 14-OCT-1999; 990S-0159330
PR 14-OCT-1999; 990S-0159331
PR 14-OCT-1999; 990S-0159637
PR 14-OCT-1999; 990S-0159638
PR 18-OCT-1999; 990S-0159584
PR 21-OCT-1999; 990S-0160741
PR 21-OCT-1999; 990S-0160757
PR 21-OCT-1999; 990S-0160768
PR 21-OCT-1999; 990S-0160770
PR 21-OCT-1999; 990S-0160814
PR 21-OCT-1999; 990S-0160815
PR 22-OCT-1999; 990S-0160880
PR 22-OCT-1999; 990S-0160981
PR 22-OCT-1999; 990S-0160989
PR 25-OCT-1999; 990S-0161404
PR 25-OCT-1999; 990S-0161405
PR 25-OCT-1999; 990S-0161406
PR 25-OCT-1999; 990S-0161359
PR 25-OCT-1999; 990S-0161360
PR 25-OCT-1999; 990S-0161361
PR 28-OCT-1999; 990S-0161920
PR 28-OCT-1999; 990S-0161992
PR 28-OCT-1999; 990S-0161993
PR 29-OCT-1999; 990S-0162142

Query Match 17.00, Score 116.5, DB CL, Length 105,
Best Local Similarity 27.88, Pred. No. 9,200,19,
Matches 35; Conservative 17; Mismatches 51; Indels 23; Gaps 5;

QY 3 AVAVFFACV-----LFCMVHKAALADIKTCNPHEFVGLVITLIGSK 46
   | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 993 asvqftctvcpkptkllwslatlllqfssalvcdkskcpa vnpfkrccp 100
   | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 47 QFEFVNLIINL-YCAOSNVKVCSDGLHTEP(DPHI)RPLSDGTNNCNVCAPLSHATL 105
   | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1050 --frccmkpccpctlnrlkqgf-pqstcdptllrllssssmccccctllsgmpt 1100
   | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 106 VAKKYA 111
   | | | |
Db 1107 lsnys 1112

RESULT 4
AA042152
ID AA042152 standard; Protein: 1193 AA.
XX
AC AA042152;
XX
LT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2000000000
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
CS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439
XX
XX 25-FEB-1999; 990S-0171826
XX 05-MAR-1999; 990S-0123180
XX 09-MAR-1999; 990S-0123548.
990S-0125788.
990S-0126264.
990S-0126785.
990S-0127462.
990S-0128244.
990S-0128714.
990S-0129845.
990S-0130077.
990S-0130449.
990S-0130510.
990S-0130891.
990S-0131449.
990S-0132048.
990S-0132407.
990S-0132484.
990S-0132485.
990S-0132486.
990S-0132487.
990S-0132863.
990S-0134256.
990S-0134218.
990S-0134219.
990S-0134221.
990S-0134370.
990S-0134768.
990S-0134941.
990S-0135124.
990S-0135353.
990S-0135629.
990S-0136321.
990S-0136392.
990S-0136782.
990S-0137222.
990S-0137528.
990S-0137502.
990S-0137724.
990S-0138094.
990S-0138540.
990S-0138847.
990S-0139119.
990S-0139452.
990S-0139453.
990S-0139492.
990S-0139454.
990S-0139455.
990S-0139456.
990S-0139457.
990S-0139458.
990S-0139459.
990S-0139460.
990S-0139461.
990S-0139462.
990S-0139463.
990S-0139750.
990S-0139763.
990S-0139817.
990S-0139899.
990S-0140353.
990S-0140354.
990S-0140695.
990S-0140823.
990S-0140991.
990S-0141287.
990S-0141842.
990S-0142154.
990S-0142055.
990S-0142390.
990S-0142803.
990S-0142920.
990S-0142977.
990S-0143542.
990S-0143624.
990S-0144005.

```

PR 16-JUL-1999;	990S-0144085;	PR 07-OCT-1999;	990S-0158029;
PR 16-JUL-1999;	990S-0144086;	PR 08-OCT-1999;	990S-0158232;
PR 19-JUL-1999;	990S-0144321;	PR 12-OCT-1999;	990S-0158469;
PR 19-JUL-1999;	990S-0144322;	PR 13-OCT-1999;	990S-0158293;
PR 19-JUL-1999;	990S-0144332;	PR 13-OCT-1999;	990S-0158294;
PR 19-JUL-1999;	990S-0144333;	PR 13-OCT-1999;	990S-0158295;
PR 19-JUL-1999;	990S-0144334;	PR 14-OCT-1999;	990S-0158329;
PR 19-JUL-1999;	990S-0144335;	PR 14-OCT-1999;	990S-0158330;
PR 20-JUL-1999;	990S-0144352;	PR 14-OCT-1999;	990S-0158331;
PR 20-JUL-1999;	990S-0144642;	PR 14-OCT-1999;	990S-0158637;
PR 20-JUL-1999;	990S-0144884;	PR 14-OCT-1999;	990S-0158638;
PR 21-JUL-1999;	990S-0144814;	PR 18-OCT-1999;	990S-0158584;
PR 21-JUL-1999;	990S-0145086;	PR 21-OCT-1999;	990S-0160741;
PR 21-JUL-1999;	990S-0145088;	PR 21-OCT-1999;	990S-0160767;
PR 22-JUL-1999;	990S-0145085;	PR 21-OCT-1999;	990S-0160768;
PR 22-JUL-1999;	990S-0145087;	PR 21-OCT-1999;	990S-0160770;
PR 22-JUL-1999;	990S-0145089;	PR 21-OCT-1999;	990S-0160814;
PR 22-JUL-1999;	990S-0145192;	PR 21-OCT-1999;	990S-0160815;
PR 23-JUL-1999;	990S-0145145;	PR 22-OCT-1999;	990S-0160980;
PR 23-JUL-1999;	990S-0145219;	PR 22-OCT-1999;	990S-0160981;
PR 23-JUL-1999;	990S-0145234;	PR 22-OCT-1999;	990S-0160989;
PR 23-JUL-1999;	990S-0145276;	PR 25-OCT-1999;	990S-0161404;
PR 27-JUL-1999;	990S-0145913;	PR 25-OCT-1999;	990S-0161405;
PR 27-JUL-1999;	990S-0145918;	PR 25-OCT-1999;	990S-0161406;
PR 27-JUL-1999;	990S-0145919;	PR 26-OCT-1999;	990S-0161359;
PR 28-JUL-1999;	990S-0145951;	PR 26-OCT-1999;	990S-0161360;
PR 02-AUG-1999;	990S-0146386;	PR 26-OCT-1999;	990S-0161361;
PR 02-AUG-1999;	990S-0146388;	PR 28-OCT-1999;	990S-0161920;
PR 02-AUG-1999;	990S-0146389;	PR 28-OCT-1999;	990S-0161992;
PR 03-AUG-1999;	990S-0146388;	PR 28-OCT-1999;	990S-0161993;
PR 03-AUG-1999;	990S-0147038;	PR 29-OCT-1999;	990S-0162142;
PR 04-AUG-1999;	990S-0147204;		
PR 04-AUG-1999;	990S-0147302;		
PR 05-AUG-1999;	990S-0147192;		
PR 05-AUG-1999;	990S-0147260;		
PR 06-AUG-1999;	990S-0147303;		
PR 06-AUG-1999;	990S-0147416;		
PR 09-AUG-1999;	990S-0147493;		
PR 09-AUG-1999;	990S-0147945;		
PR 10-AUG-1999;	990S-0148171;		
PR 11-AUG-1999;	990S-0148339;		
PR 12-AUG-1999;	990S-0148441;		
PR 13-AUG-1999;	990S-0148565;		
PR 14-AUG-1999;	990S-0148684;		
PR 16-AUG-1999;	990S-0149368;		
PR 17-AUG-1999;	990S-0149175;		
PR 18-AUG-1999;	990S-0149426;		
PR 20-AUG-1999;	990S-0149722;		
PR 20-AUG-1999;	990S-0149723;		
PR 20-AUG-1999;	990S-0149929;		
PR 23-AUG-1999;	990S-0149902;		
PR 23-AUG-1999;	990S-0149930;		
PR 25-AUG-1999;	990S-0150566;		
PR 26-AUG-1999;	990S-0150884;		
PR 27-AUG-1999;	990S-0151055;		
PR 27-AUG-1999;	990S-0151066;		
PR 27-AUG-1999;	990S-0151080;		
PR 30-AUG-1999;	990S-0151303;		
PR 31-AUG-1999;	990S-0151448;		
PR 01-SEP-1999;	990S-0151940;		
PR 07-SEP-1999;	990S-0152363;		
PR 10-SEP-1999;	990S-0153070;		
PR 13-SEP-1999;	990S-0153758;		
PR 15-SEP-1999;	990S-0154018;		
PR 16-SEP-1999;	990S-0154039;		
PR 20-SEP-1999;	990S-0154779;		
PR 22-SEP-1999;	990S-0155149;		
PR 23-SEP-1999;	990S-0155486;		
PR 24-SEP-1999;	990S-0155659;		
PR 28-SEP-1999;	990S-0156458;		
PR 29-SEP-1999;	990S-0156596;		
PR 04-OCT-1999;	990S-0157117;		
PR 05-OCT-1999;	990S-0157753;		
PR 06-OCT-1999;	990S-0157865;		
Query Match			
Best Local Similarity 17.08; Score 116.5; EB 21; Length 1194;			
Matches 35; Conservative 17; Mismatches 51; Indels 23; Gaps 5;			
QY 3 AVAVFFACV	-----LFTMVHKAALADKTNPEDEWVQITGLTIQK 46		
PR 105 AAVVPTGCPKPKVETLWSELAIIIPSTVQVQKPKSQA VRTPKKTPPK 1113			
QY 47 QLETVNLNNL YFAQSNVKSQDLHITFDQPHIRPLEDQINRGLVNGALISHATL 105			
DB 1114 --frrvrmvkvpmvptlnal:146qtf:fpstlydptfrrvssaqvavvnddtpispmqt 1170			
QY 106 VAPKVA 111			
DB 1171 Istnys 1176			
RESULT 5			
AAG11089			
ID AAG11089 standard; Protein: 135 AA.			
XX			
AC AAG11089;			
XX			
DI 17-OCT-2000 (first entry)			
XX			
DE			
XX			
XX			
KW			
KW			
XX			
OS			
XX			
PN			
XX			
PD			
XX			
PF			
XX			
PR			
PR			

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence;

Arabidopsis thaliana.

EP1334305-A2.

06-SEP-2000.

25-FEB 2000; 2000EP 0401439.

25-FEB 1999; 990S 0121825.

05-MAR-1999; 990S 0121800.

PP 09-MAR-1999; 99US-0123549;
PP 23-MAR-1999; 99US-0125788;
PR 25-MAR-1999; 99US-0126264;
PP 25-MAR-1999; 99US-0126785;
PR 01-APR-1999; 99US-0127462;
PR 06-APR-1999; 99US-0128234;
PR 08-APR-1999; 99US-0128714;
PR 16-APR-1999; 99US-0129845;
PR 19-APR-1999; 99US-0130077;
PR 21-APR-1999; 99US-0130449;
PR 23-APR-1999; 99US-0130510;
PR 23-APR-1999; 99US-0130891;
PR 28-APR-1999; 99US-0131449;
PR 30-APR-1999; 99US-0132048;
PR 30-APR-1999; 99US-0132407;
PR 04-MAY-1999; 99US-0132484;
PR 05-MAY-1999; 99US-0132485;
PR 06-MAY-1999; 99US-0132486;
PR 06-MAY-1999; 99US-0132487;
PR 07-MAY-1999; 99US-0132863;
PR 11-MAY-1999; 99US-0134256;
PR 14-MAY-1999; 99US-0134218;
PR 14-MAY-1999; 99US-0134219;
PR 14-MAY-1999; 99US-0134221;
PR 14-MAY-1999; 99US-0134370;
PR 18-MAY-1999; 99US-0134768;
PR 19-MAY-1999; 99US-0134941;
PR 20-MAY-1999; 99US-0135124;
PR 21-MAY-1999; 99US-0135353;
PR 24-MAY-1999; 99US-0135629;
PR 25-MAY-1999; 99US-0136021;
PR 27-MAY-1999; 99US-0136392;
PR 28-MAY-1999; 99US-0136782;
PR 01-JUN-1999; 99US-0137222;
PR 03-JUN-1999; 99US-0137528;
PR 04-JUN-1999; 99US-0137502;
PR 07-JUN-1999; 99US-0137724;
PR 08-JUN-1999; 99US-0138044;
PR 10-JUN-1999; 99US-0138540;
PR 10-JUN-1999; 99US-0138847;
PR 14-JUN-1999; 99US-0139119;
PR 16-JUN-1999; 99US-0139455;
PR 16-JUN-1999; 99US-0139453;
PR 17-JUN-1999; 99US-0139492;
PR 18-JUN-1999; 99US-0139454;
PR 18-JUN-1999; 99US-0139455;
PR 18-JUN-1999; 99US-0139456;
PR 18-JUN-1999; 99US-0139457;
PR 18-JUN-1999; 99US-0139458;
PR 18-JUN-1999; 99US-0139459;
PR 18-JUN-1999; 99US-0139460;
PR 18-JUN-1999; 99US-0139461;
PR 18-JUN-1999; 99US-0139462;
PR 18-JUN-1999; 99US-0140454;
PR 18-JUN-1999; 99US-0140465;
PR 28-JUN-1999; 99US-0140823;
PR 29-JUN-1999; 99US-0140991;
PR 30-JUN-1999; 99US-0141287;
PR 01-JUL-1999; 99US-0141842;
PR 01-JUL-1999; 99US-0142154;
PR 02-JUL-1999; 99US-0142065;
PR 06-JUL-1999; 99US-0142390;
PR 08-JUL-1999; 99US-0142803;
PR 09-JUL-1999; 99US-0142920;
PR 12-JUL-1999; 99US-0142977;
PR 13-JUL-1999; 99US-0143542;
PR 14-JUL-1999; 99US-0143624;
PP 15-JUL-1999; 99US-0144005;
PR 16-JUL-1999; 99US-0144085;
PP 16-JUL-1999; 99US-0144086;
PR 19-JUL-1999; 99US-0144325;
PR 19-JUL-1999; 99US-0144331;
PR 19-JUL-1999; 99US-0144332;
PR 19-JUL-1999; 99US-0144333;
PR 19-JUL-1999; 99US-0144334;
PR 19-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144352;
PR 20-JUL-1999; 99US-0144362;
PR 20-JUL-1999; 99US-0144884;
PR 21-JUL-1999; 99US-0144814;
PR 21-JUL-1999; 99US-0145086;
PR 21-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145085;
PR 22-JUL-1999; 99US-0145087;
PR 22-JUL-1999; 99US-0145089;
PR 22-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145145;
PR 23-JUL-1999; 99US-0145218;
PR 24-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 27-JUL-1999; 99US-0145919;
PR 02-AUG-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146366;
PR 02-AUG-1999; 99US-0146388;
PR 02-AUG-1999; 99US-0146389;
PR 02-AUG-1999; 99US-0147038;
PR 03-AUG-1999; 99US-0147204;
PR 03-AUG-1999; 99US-0147202;
PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 06-AUG-1999; 99US-0147493;
PR 06-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 24-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 28-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;

PR 25-FEB-1999; 9905-0121825;
PR 05-MAR-1999; 9905-0123180;
PR 09-MAR-1999; 9905-0123548;
PR 23-MAR-1999; 9905-0125788;
PR 25-MAR-1999; 9905-0126264;
PR 29-MAR-1999; 9905-0126785;
PR 01-APR-1999; 9905-0127442;
PR 06-APR-1999; 9905-0128244;
PR 08-APR-1999; 9905-0128714;
PR 16-APR-1999; 9905-0129845;
PR 19-APR-1999; 9905-0130077;
PR 21-APR-1999; 9905-0130449;
PR 24-APR-1999; 9905-0130510;
PR 23-APR-1999; 9905-0130891;
PR 28-APR-1999; 9905-0131449;
PR 30-APR-1999; 9905-0132048;
PR 30-APR-1999; 9905-0132407;
PR 04-MAY-1999; 9905-0132484;
PR 05-MAY-1999; 9905-0132485;
PR 06-MAY-1999; 9905-0132486;
PR 07-MAY-1999; 9905-0132487;
PR 11-MAY-1999; 9905-0132863;
PR 11-MAY-1999; 9905-0134256;
PR 14-MAY-1999; 9905-0134218;
PR 14-MAY-1999; 9905-0134219;
PR 14-MAY-1999; 9905-0134221;
PR 14-MAY-1999; 9905-0134370;
PR 18-MAY-1999; 9905-0134768;
PR 19-MAY-1999; 9905-0134941;
PR 20-MAY-1999; 9905-0135124;
PR 21-MAY-1999; 9905-0135353;
PR 24-MAY-1999; 9905-0135626;
PR 25-MAY-1999; 9905-0136021;
PR 27-MAY-1999; 9905-0136392;
PR 28-MAY-1999; 9905-0136782;
PR 01-JUN-1999; 9905-0137222;
PR 03-JUN-1999; 9905-0137528;
PR 04-JUN-1999; 9905-0137502;
PR 07-JUN-1999; 9905-0137724;
PR 08-JUN-1999; 9905-0138094;
PR 10-JUN-1999; 9905-0138540;
PR 10-JUN-1999; 9905-0138847;
PR 14-JUN-1999; 9905-0139119;
PR 16-JUN-1999; 9905-0139452;
PR 16-JUN-1999; 9905-0139453;
PR 17-JUN-1999; 9905-0139459;
PR 18-JUN-1999; 9905-0139454;
PR 18-JUN-1999; 9905-0139455;
PR 18-JUN-1999; 9905-0139456;
PR 18-JUN-1999; 9905-0139457;
PR 18-JUN-1999; 9905-0139458;
PR 18-JUN-1999; 9905-0139459;
PR 18-JUN-1999; 9905-0139460;
PR 18-JUN-1999; 9905-0139461;
PR 18-JUN-1999; 9905-0139462;
PR 18-JUN-1999; 9905-0139463;
PR 18-JUN-1999; 9905-0139760;
PR 18-JUN-1999; 9905-0139763;
PR 21-JUN-1999; 9905-0139817;
PR 22-JUN-1999; 9905-0139899;
PR 24-JUN-1999; 9905-0140353;
PR 24-JUN-1999; 9905-0140354;
PR 24-JUN-1999; 9905-0140695;
PR 28-JUN-1999; 9905-0140823;
PR 29-JUN-1999; 9905-0140991;
PR 30-JUN-1999; 9905-0141287;
PR 01-JUL-1999; 9905-0141842;
PR 01-JUL-1999; 9905-0142154;
PR 02-JUL-1999; 9905-0142055;
PR 06-JUL-1999; 9905-0142390;
PR 08-JUL-1999; 9905-0142863;
PR 09-JUL-1999; 9905-0142920;
PR 12-JUL-1999; 9905-0142977;
PR 13-JUL-1999; 9905-0143542;
PR 14-JUL-1999; 9905-0143624;
PR 15-JUL-1999; 9905-0144005;
PR 16-JUL-1999; 9905-0144085;
PR 16-JUL-1999; 9905-0144086;
PR 19-JUL-1999; 9905-0144325;
PR 19-JUL-1999; 9905-0144341;
PR 19-JUL-1999; 9905-0144342;
PR 19-JUL-1999; 9905-0144343;
PR 19-JUL-1999; 9905-0144344;
PR 19-JUL-1999; 9905-0144345;
PR 20-JUL-1999; 9905-0144352;
PR 20-JUL-1999; 9905-0144432;
PR 20-JUL-1999; 9905-0144884;
PR 21-JUL-1999; 9905-0144814;
PR 21-JUL-1999; 9905-0145086;
PR 21-JUL-1999; 9905-0145088;
PR 22-JUL-1999; 9905-0145085;
PR 22-JUL-1999; 9905-0145087;
PR 22-JUL-1999; 9905-0145089;
PR 22-JUL-1999; 9905-0145102;
PR 23-JUL-1999; 9905-0145145;
PR 23-JUL-1999; 9905-0145218;
PR 23-JUL-1999; 9905-0145224;
PR 26-JUL-1999; 9905-0145276;
PR 27-JUL-1999; 9905-0145913;
PR 27-JUL-1999; 9905-0145918;
PR 27-JUL-1999; 9905-0145919;
PR 28-JUL-1999; 9905-0145951;
PR 02-AUG-1999; 9905-0146386;
PR 02-AUG-1999; 9905-0146388;
PR 03-AUG-1999; 9905-0147038;
PR 04-AUG-1999; 9905-0147204;
PR 04-AUG-1999; 9905-0147302;
PR 05-AUG-1999; 9905-0147342;
PR 05-AUG-1999; 9905-0147260;
PR 06-AUG-1999; 9905-0147304;
PR 06-AUG-1999; 9905-0147416;
PR 09-AUG-1999; 9905-0147493;
PR 09-AUG-1999; 9905-0147945;
PR 10-AUG-1999; 9905-0148171;
PR 11-AUG-1999; 9905-0148319;
PR 12-AUG-1999; 9905-0148341;
PR 13-AUG-1999; 9905-0148565;
PR 13-AUG-1999; 9905-0148684;
PR 16-AUG-1999; 9905-0149368;
PR 17-AUG-1999; 9905-0149175;
PR 18-AUG-1999; 9905-0149426;
PR 20-AUG-1999; 9905-0149722;
PR 20-AUG-1999; 9905-0149723;
PR 20-AUG-1999; 9905-0149924;
PR 24-AUG-1999; 9905-0149902;
PR 24-AUG-1999; 9905-0149930;
PR 25-AUG-1999; 9905-0150566;
PR 26-AUG-1999; 9905-0150884;
PR 27-AUG-1999; 9905-0151066;
PR 27-AUG-1999; 9905-0151080;
PR 30-AUG-1999; 9905-0151303;
PR 31-AUG-1999; 9905-0151438;
PR 01-SEP-1999; 9905-0151930;
PR 07-SEP-1999; 9905-0152363;
PR 10-SEP-1999; 9905-0153070;
PR 13-SEP-1999; 9905-0153758;
PR 15-SEP-1999; 9905-0154018;
PR 16-SEP-1999; 9905-0154039;
PR 20-SEP-1999; 9905-0154779;
PR 22-SEP-1999; 9905-0155139;
PR 23-SEP-1999; 9905-0155486;
PR 24-SEP-1999; 9905-0155654;
PR 28-SEP-1999; 9905-0156458;
PR 29-SEP-1999; 9905-0156596;

PR 16-JUL-1999; 990S-0144085.
 PR 16-JUL-1999; 990S-0144086.
 PR 19-JUL-1999; 990S-0144327.
 PR 19-JUL-1999; 990S-0144341.
 PR 19-JUL-1999; 990S-0144342.
 PR 19-JUL-1999; 990S-0144343.
 PR 19-JUL-1999; 990S-0144344.
 PR 19-JUL-1999; 990S-0144345.
 PR 20-JUL-1999; 990S-0144352.
 PR 20-JUL-1999; 990S-0144642.
 PR 20-JUL-1999; 990S-0144894.
 PR 21-JUL-1999; 990S-0144814.
 PR 21-JUL-1999; 990S-0145086.
 PR 21-JUL-1999; 990S-0145088.
 PR 22-JUL-1999; 990S-0145045.
 PR 22-JUL-1999; 990S-0145047.
 PR 22-JUL-1999; 990S-0145049.
 PR 22-JUL-1999; 990S-0145091.
 PR 23-JUL-1999; 990S-0145145.
 PR 23-JUL-1999; 990S-0145218.
 PR 24-JUL-1999; 990S-0145224.
 PR 26-JUL-1999; 990S-0145276.
 PR 27-JUL-1999; 990S-0145918.
 PR 27-JUL-1999; 990S-0145918.
 PR 27-JUL-1999; 990S-0145919.
 PR 28-JUL-1999; 990S-0145951.
 PR 02-AUG-1999; 990S-0146386.
 PR 02-AUG-1999; 990S-0146388.
 PR 02-AUG-1999; 990S-0146389.
 PR 03-AUG-1999; 990S-0147038.
 PR 04-AUG-1999; 990S-0147204.
 PR 04-AUG-1999; 990S-0147302.
 PR 05-AUG-1999; 990S-0147192.
 PR 05-AUG-1999; 990S-0147260.
 PR 06-AUG-1999; 990S-0147303.
 PR 06-AUG-1999; 990S-0147416.
 PR 09-AUG-1999; 990S-0147493.
 PR 09-AUG-1999; 990S-0147935.
 PR 10-AUG-1999; 990S-0148171.
 PR 11-AUG-1999; 990S-0148319.
 PR 12-AUG-1999; 990S-0148341.
 PR 13-AUG-1999; 990S-0148565.
 PR 13-AUG-1999; 990S-0148684.
 PR 16-AUG-1999; 990S-0149368.
 PR 17-AUG-1999; 990S-0149176.
 PR 18-AUG-1999; 990S-0149426.
 PR 20-AUG-1999; 990S-0149722.
 PR 20-AUG-1999; 990S-0149723.
 PR 20-AUG-1999; 990S-0149929.
 PR 23-AUG-1999; 990S-0149902.
 PR 23-AUG-1999; 990S-0149930.
 PR 25-AUG-1999; 990S-0150566.
 PR 26-AUG-1999; 990S-0150884.
 PR 27-AUG-1999; 990S-0151065.
 PR 27-AUG-1999; 990S-0151066.
 PR 27-AUG-1999; 990S-0151080.
 PR 30-AUG-1999; 990S-0151303.
 PR 31-AUG-1999; 990S-0151438.
 PR 01-SEP-1999; 990S-0151930.
 PR 07-SEP-1999; 990S-0152463.
 PR 10-SEP-1999; 990S-0153070.
 PR 13-SEP-1999; 990S-0153758.
 PR 15-SEP-1999; 990S-0154018.
 PR 16-SEP-1999; 990S-0154034.
 PR 20-SEP-1999; 990S-0154774.
 PR 22-SEP-1999; 990S-0155119.
 PR 23-SEP-1999; 990S-0155486.
 PR 24-SEP-1999; 990S-0155486.
 PR 28-SEP-1999; 990S-0156458.
 PR 29-SEP-1999; 990S-0156596.
 PR 04-OCT-1999; 990S-0157117.
 PR 05-OCT-1999; 990S-0157753.
 PR 06-OCT-1999; 990S-0157845.

PR 07-OCT-1999; 990S-0158029.
 PR 08-OCT-1999; 990S-0158232.
 PR 12-OCT-1999; 990S-0158369.
 PR 13-OCT-1999; 990S-0159293.
 PR 13-OCT-1999; 990S-0159294.
 PR 13-OCT-1999; 990S-0159295.
 PR 14-OCT-1999; 990S-0159329.
 PR 14-OCT-1999; 990S-0159330.
 PR 14-OCT-1999; 990S-0159331.
 PR 14-OCT-1999; 990S-0159637.
 PR 14-OCT-1999; 990S-0159638.
 PR 18-OCT-1999; 990S-0159584.
 PR 21-OCT-1999; 990S-0160741.
 PR 21-OCT-1999; 990S-0160767.
 PR 21-OCT-1999; 990S-0160768.
 PR 21-OCT-1999; 990S-0160770.
 PR 21-OCT-1999; 990S-0160814.
 PR 21-OCT-1999; 990S-0160815.
 PR 22-OCT-1999; 990S-0160980.
 PR 22-OCT-1999; 990S-0160981.
 PR 22-OCT-1999; 990S-0160989.
 PR 25-OCT-1999; 990S-0161404.
 PR 25-OCT-1999; 990S-0161405.
 PR 25-OCT-1999; 990S-0161406.
 PR 26-OCT-1999; 990S-0161359.
 PR 26-OCT-1999; 990S-0161360.
 PR 26-OCT-1999; 990S-0161361.
 PR 28-OCT-1999; 990S-0161920.
 PR 28-OCT-1999; 990S-0161992.
 PR 28-OCT-1999; 990S-0161993.
 PR 29-OCT-1999; 990S-0162142.

Query Match 13.98; Score 95.5; DN 21; Length 70;

Best Local Similarity 34.08; Pred. No. 0.0011;

Matches 18; Conservative 11; Mismatches 23; Indels 1; Gaps 1;

QY 59 CAQSNVKSQGLHTPEIDPHIPLSLDGTNNCLVNGAPISHATLVAAPKVA L11

DB 2 cpilnlrikcppt-pqslpdtflrissssanccvznadlplspmqtlstnys 53

RESULT 9

AAG11091

ID AAG11091 standard; Protein: 75 AA.

AC AAG11091;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9677.

XX

KW Protein identification, signal transduction pathway, metabolic pathway;

KW hybridisation assay, genetic mapping, gene expression control; promoter;

KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1034405.A2.

XX

PD 06-SEP-2000.

XX

PF 25-FER-2000; 2000EP-0301439.

XX

PR 25-FER-1999; 990S-0121825.

PR 05-MAR-1999; 990S-0123180.

PR 09-MAR-1999; 990S-0123548.

PR 25-MAR-1999; 990S-0125788.

PR 29-MAR-1999; 990S-0126264.

PR 01-APR-1999; 990S-0126785.

PR 06-APR-1999; 990S-0127462.

PR 08-APR-1999; 990S-0128234.

PR 16-APR-1999; 990S-0128714.

PR 16-APR-1999; 990S-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130444
PR 23-APR-1999; 99US-0130510
PR 25-APR-1999; 99US-0130801
PR 28-APR-1999; 99US-0131419
PR 30-APR-1999; 99US-0132048
PR 30-APR-1999; 99US-0132407
PR 04-MAY-1999; 99US-0132484
PR 05-MAY-1999; 99US-0132485
PR 06-MAY-1999; 99US-0132486
PR 06-MAY-1999; 99US-0132487
PR 07-MAY-1999; 99US-0132863
PR 11-MAY-1999; 99US-0134256
PR 14-MAY-1999; 99US-0134218
PR 14-MAY-1999; 99US-0134219
PR 14-MAY-1999; 99US-0134221
PR 14-MAY-1999; 99US-0134370
PR 18-MAY-1999; 99US-0134768
PR 19-MAY-1999; 99US-0134941
PR 20-MAY-1999; 99US-0135124
PR 21-MAY-1999; 99US-0135253
PR 24-MAY-1999; 99US-0135629
PR 25-MAY-1999; 99US-0136021
PR 27-MAY-1999; 99US-0136392
PR 28-MAY-1999; 99US-0136782
PR 01-JUN-1999; 99US-0137222
PR 03-JUN-1999; 99US-0137528
PR 04-JUN-1999; 99US-0137502
PR 07-JUN-1999; 99US-0137724
PR 08-JUN-1999; 99US-0138054
PR 10-JUN-1999; 99US-0138540
PR 10-JUN-1999; 99US-0138847
PR 14-JUN-1999; 99US-0139119
PR 16-JUN-1999; 99US-0139452
PR 16-JUN-1999; 99US-0139453
PR 17-JUN-1999; 99US-0139492
PR 18-JUN-1999; 99US-0139454
PR 18-JUN-1999; 99US-0139455
PR 18-JUN-1999; 99US-0139456
PR 18-JUN-1999; 99US-0139457
PR 18-JUN-1999; 99US-0139458
PR 18-JUN-1999; 99US-0139459
PR 18-JUN-1999; 99US-0139460
PR 18-JUN-1999; 99US-0139461
PR 18-JUN-1999; 99US-0139462
PR 18-JUN-1999; 99US-0139463
PR 18-JUN-1999; 99US-0139750
PR 18-JUN-1999; 99US-0139764
PR 21-JUN-1999; 99US-0139817
PR 22-JUN-1999; 99US-0140044
PR 23-JUN-1999; 99US-0140353
PR 24-JUN-1999; 99US-0140454
PR 24-JUN-1999; 99US-0140695
PR 26-JUN-1999; 99US-0140823
PR 26-JUN-1999; 99US-0140941
PR 30-JUN-1999; 99US-0141287
PR 01-JUL-1999; 99US-0141842
PR 01-JUL-1999; 99US-0142154
PR 02-JUL-1999; 99US-0142055
PR 06-JUL-1999; 99US-0142390
PR 08-JUL-1999; 99US-0142903
PR 09-JUL-1999; 99US-0142920
PR 12-JUL-1999; 99US-0142977
PR 13-JUL-1999; 99US-0143542
PR 14-JUL-1999; 99US-0143624
PR 15-JUL-1999; 99US-0144005
PR 16-JUL-1999; 99US-0144086
PR 16-JUL-1999; 99US-0144325
PR 19-JUL-1999; 99US-0144331
PR 19-JUL-1999; 99US-0144332
PR 19-JUL-1999; 99US-0144333
PR 19-JUL-1999; 99US-0144334
PR 19-JUL-1999; 99US-0144335
PR 20-JUL-1999; 99US-0144352
PR 20-JUL-1999; 99US-0144353
PR 20-JUL-1999; 99US-0144884
PR 21-JUL-1999; 99US-0144814
PR 21-JUL-1999; 99US-0145086
PR 21-JUL-1999; 99US-0145088
PR 22-JUL-1999; 99US-0145085
PR 22-JUL-1999; 99US-0145087
PR 22-JUL-1999; 99US-0145089
PR 23-JUL-1999; 99US-0145192
PR 23-JUL-1999; 99US-0145145
PR 23-JUL-1999; 99US-0145218
PR 23-JUL-1999; 99US-0145224
PR 26-JUL-1999; 99US-0145276
PR 27-JUL-1999; 99US-0145913
PR 27-JUL-1999; 99US-0145918
PR 27-JUL-1999; 99US-0145919
PR 28-JUL-1999; 99US-0145951
PR 28-JUL-1999; 99US-0146386
PR 02-AUG-1999; 99US-0146388
PR 02-AUG-1999; 99US-0146389
PR 03-AUG-1999; 99US-0147038
PR 04-AUG-1999; 99US-0147204
PR 04-AUG-1999; 99US-0147302
PR 05-AUG-1999; 99US-0147192
PR 05-AUG-1999; 99US-0147260
PR 06-AUG-1999; 99US-0147303
PR 06-AUG-1999; 99US-0147416
PR 09-AUG-1999; 99US-0147493
PR 09-AUG-1999; 99US-0147935
PR 10-AUG-1999; 99US-0148171
PR 11-AUG-1999; 99US-0148319
PR 12-AUG-1999; 99US-0148341
PR 13-AUG-1999; 99US-0148565
PR 13-AUG-1999; 99US-0148684
PR 15-AUG-1999; 99US-0149368
PR 17-AUG-1999; 99US-0149175
PR 18-AUG-1999; 99US-0149426
PR 20-AUG-1999; 99US-0149723
PR 20-AUG-1999; 99US-0149723
PR 20-AUG-1999; 99US-0149929
PR 23-AUG-1999; 99US-0149602
PR 23-AUG-1999; 99US-0149930
PR 25-AUG-1999; 99US-0150566
PR 25-AUG-1999; 99US-0150884
PR 27-AUG-1999; 99US-0151065
PR 27-AUG-1999; 99US-0151866
PR 27-AUG-1999; 99US-0151080
PR 30-AUG-1999; 99US-0151303
PR 31-AUG-1999; 99US-0151338
PR 01-SEP-1999; 99US-0151949
PR 07-SEP-1999; 99US-0152363
PR 10-SEP-1999; 99US-0154070
PR 12-SEP-1999; 99US-0153758
PR 15-SEP-1999; 99US-0154018
PR 16-SEP-1999; 99US-0154034
PR 20-SEP-1999; 99US-0154779
PR 22-SEP-1999; 99US-0155139
PR 23-SEP-1999; 99US-0155486
PR 24-SEP-1999; 99US-0155659
PR 28-SEP-1999; 99US-0156458
PR 29-SEP-1999; 99US-0156596
PR 04-OCT-1999; 99US-0157117
PR 05-OCT-1999; 99US-0157753
PR 06-OCT-1999; 99US-0157865
PR 07-OCT-1999; 99US-0159029
PR 08-OCT-1999; 99US-0159232
PR 12-OCT-1999; 99US-0159369
PR 13-OCT-1999; 99US-0159293
PR 13-OCT-1999; 99US-0159294
PR 13-OCT-1999; 99US-0159295
PR 13-OCT-1999; 99US-0159329
PR 13-OCT-1999; 99US-0159330

PR 04-JUN-1999;	990S-0137502	PR 06-AUG-1999;	990S-0147303
PR 07-JUN-1999;	990S-0137724	PR 06-AUG-1999;	990S-0147416
PR 08-JUN-1999;	990S-0138094	PR 09-AUG-1999;	990S-0147493
PR 10-JUN-1999;	990S-0138540	PR 09-AUG-1999;	990S-0147935
PR 10-JUN-1999;	990S-0138847	PR 10-AUG-1999;	990S-0148171
PR 14-JUN-1999;	990S-0139119	PR 11-AUG-1999;	990S-0148319
PR 16-JUN-1999;	990S-0139453	PR 12-AUG-1999;	990S-0148341
PR 16-JUN-1999;	990S-0139453	PR 13-AUG-1999;	990S-0148565
PR 17-JUN-1999;	990S-0139452	PR 13-AUG-1999;	990S-0148684
PR 18-JUN-1999;	990S-0139454	PR 16-AUG-1999;	990S-0149368
PR 18-JUN-1999;	990S-0139455	PR 17-AUG-1999;	990S-0149175
PR 18-JUN-1999;	990S-0139456	PR 18-AUG-1999;	990S-0149426
PR 18-JUN-1999;	990S-0139457	PR 20-AUG-1999;	990S-0149722
PR 18-JUN-1999;	990S-0139458	PR 20-AUG-1999;	990S-0149723
PR 18-JUN-1999;	990S-0139459	PR 20-AUG-1999;	990S-0149929
PR 18-JUN-1999;	990S-0139460	PR 23-AUG-1999;	990S-0149902
PR 18-JUN-1999;	990S-0139461	PR 25-AUG-1999;	990S-0149930
PR 18-JUN-1999;	990S-0139462	PR 26-AUG-1999;	990S-0150566
PR 18-JUN-1999;	990S-0139463	PR 26-AUG-1999;	990S-0150884
PR 18-JUN-1999;	990S-0139750	PR 27-AUG-1999;	990S-0151065
PR 18-JUN-1999;	990S-0139763	PR 27-AUG-1999;	990S-0151066
PR 21-JUN-1999;	990S-0139817	PR 27-AUG-1999;	990S-0151080
PR 22-JUN-1999;	990S-0139890	PR 30-AUG-1999;	990S-0151303
PR 23-JUN-1999;	990S-0140353	PR 31-AUG-1999;	990S-0151438
PR 23-JUN-1999;	990S-0140354	PR 01-SEP-1999;	990S-0151930
PR 24-JUN-1999;	990S-0140695	PR 07-SEP-1999;	990S-0152363
PR 28-JUN-1999;	990S-0140824	PR 10-SEP-1999;	990S-0152370
PR 29-JUN-1999;	990S-0140941	PR 14-SEP-1999;	990S-0153758
PR 30-JUN-1999;	990S-0141287	PR 15-SEP-1999;	990S-0154018
PR 01-JUL-1999;	990S-0141842	PR 16-SEP-1999;	990S-0154039
PR 01-JUL-1999;	990S-0142154	PR 20-SEP-1999;	990S-0154779
PR 02-JUL-1999;	990S-0142055	PR 22-SEP-1999;	990S-0155139
PR 06-JUL-1999;	990S-0142390	PR 23-SEP-1999;	990S-0155486
PR 08-JUL-1999;	990S-0142803	PR 24-SEP-1999;	990S-0155659
PR 09-JUL-1999;	990S-0142920	PR 28-SEP-1999;	990S-0156458
PR 12-JUL-1999;	990S-0142977	PR 29-SEP-1999;	990S-0156596
PR 13-JUL-1999;	990S-0143542	PR 04-OCT-1999;	990S-0157117
PR 14-JUL-1999;	990S-0143624	PR 05-OCT-1999;	990S-0157753
PR 15-JUL-1999;	990S-0144005	PR 06-OCT-1999;	990S-0157865
PR 16-JUL-1999;	990S-0144085	PR 07-OCT-1999;	990S-0158029
PR 16-JUL-1999;	990S-0144086	PR 08-OCT-1999;	990S-0158242
PR 19-JUL-1999;	990S-0144325	PR 12-OCT-1999;	990S-0158369
PR 19-JUL-1999;	990S-0144331	PR 13-OCT-1999;	990S-0159293
PR 19-JUL-1999;	990S-0144332	PR 13-OCT-1999;	990S-0159294
PR 19-JUL-1999;	990S-0144333	PR 13-OCT-1999;	990S-0159295
PR 19-JUL-1999;	990S-0144334	PR 14-OCT-1999;	990S-0159329
PR 19-JUL-1999;	990S-0144335	PR 14-OCT-1999;	990S-0159340
PR 20-JUL-1999;	990S-0144352	PR 14-OCT-1999;	990S-0159331
PR 20-JUL-1999;	990S-0144632	PR 14-OCT-1999;	990S-0159637
PR 21-JUL-1999;	990S-0144884	PR 14-OCT-1999;	990S-0159638
PR 21-JUL-1999;	990S-0144814	PR 18-OCT-1999;	990S-0159584
PR 21-JUL-1999;	990S-0145086	PR 21-OCT-1999;	990S-0160741
PR 21-JUL-1999;	990S-0145088	PR 21-OCT-1999;	990S-0160767
PR 22-JUL-1999;	990S-0145085	PR 21-OCT-1999;	990S-0160768
PR 22-JUL-1999;	990S-0145087	PR 21-OCT-1999;	990S-0160770
PR 22-JUL-1999;	990S-0145089	PR 21-OCT-1999;	990S-0160814
PR 22-JUL-1999;	990S-0145152	PR 21-OCT-1999;	990S-0160815
PR 23-JUL-1999;	990S-0145145	PR 22-OCT-1999;	990S-0160980
PR 23-JUL-1999;	990S-0145218	PR 22-OCT-1999;	990S-0160981
PR 23-JUL-1999;	990S-0145224	PR 22-OCT-1999;	990S-0160989
PR 26-JUL-1999;	990S-0145226	PR 25-OCT-1999;	990S-0161404
PR 27-JUL-1999;	990S-0145913	PR 25-OCT-1999;	990S-0161405
PR 27-JUL-1999;	990S-0145918	PR 25-OCT-1999;	990S-0161406
PR 27-JUL-1999;	990S-0145919	PR 26-OCT-1999;	990S-0161359
PR 28-JUL-1999;	990S-0145951	PR 26-OCT-1999;	990S-0161360
PR 02-AUG-1999;	990S-0146386	PR 26-OCT-1999;	990S-0161361
PR 02-AUG-1999;	990S-0146388	PR 28-OCT-1999;	990S-0161920
PR 03-AUG-1999;	990S-0146389	PR 28-OCT-1999;	990S-0161922
PR 04-AUG-1999;	990S-0147204	PR 28-OCT-1999;	990S-0161993
PR 04-AUG-1999;	990S-0147302	PR 29-OCT-1999;	990S-0162142
PR 05-AUG-1999;	990S-0147192		
PR 05-AUG-1999;	990S-0147260		

Query Match
Best Local Similarity

11.28;
30.58;

Score 74.5; PP 21;
Pred. No. 1.1; Length 281;


```

Matches 36; Conservative 11; Mismatches 44; Indels 27; Gaps 9.
QY 18 KAALADIKTNPDPVMTQTITGLTIGSGKDFEVIIN-NLYCAU-----SNVKVSAAG 70
DB 96 kqaladegcrvyvldvdrvg-----nfhsobhahlyvagnlfggsknvnys hm 147
QY 71 LHTE--PIDPHIPLSDGTNCLVNSGAPISHAILVAFKYAMN-VDPSPSTLSSSL 125
DE 148 italsgkykypjllqpldd turl... . hpsgthgghgghgghgghgghggh 157

RESULT 14
AAG23733
ID AAG23733 standard; Protein: 297 AA.
XX
AC AAG23733;
XX
DI 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27148.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
IN EP1033405-A2.
XX
FO 06-SEP-2000.
XX
FE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 04-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0125264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 24-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132497.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28 MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139452.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 08-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 16-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 23-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145413.
PR 27-JUL-1999; 990S-0145418.
PR 27-JUL-1999; 990S-0145419.
PR 28-JUL-1999; 990S-0145451.
PR 02-AUG-1999; 990S-0145486.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147403.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.

```


00 71 LHTTE--RHHPHIIRGSGTNNCLNNGCARTSHATLVAFKYAWD-VDPSESTLSSUI 126
10 184 ldlstapkyqrbupl34--rls-----l4rsglvyv34lpmvrylsk32 232

Search completed: July 18, 2001, 17:27:58
Job time: 40 sec

A: description: structural protein forming filaments in striated muscle
C: Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; gI
Keywords: alternative splicing; calmodulin binding; cardiac muscle; duplication; gI
structural protein
E: 21752-23081, human; F: 1016-1036, human homology - EKH
F: 94-177, 905-2378, 2478-2491, 2543-2649, 2763-2836, 3089-3179, 3394-3432, 3649-3772-4040, 4199-4299, 4399-4499, 4599-4699, 4799-4899, 4999-5099, 5199-5299, 5399-5499, 5599-5699, 5799-5899, 5999-6099, 6199-6299, 6399-6499, 6599-6699, 6799-6899, 6999-7099, 7199-7299, 7399-7499, 7599-7699, 7799-7899, 7999-8099, 8199-8299, 8399-8499, 8599-8699, 8799-8899, 8999-9099, 9199-9299, 9399-9499, 9599-9699, 9799-9899, 9999-10099, 10199-10299, 10399-10499, 10599-10699, 10799-10899, 10999-11099, 11199-11299, 11399-11499, 11599-11699, 11799-11899, 11999-12099, 12199-12299, 12399-12499, 12599-12699, 12799-12899, 12999-13099, 13199-13299, 13399-13499, 13599-13699, 13799-13899, 13999-14099, 14199-14299, 14399-14499, 14599-14699, 14799-14899, 14999-15099, 15199-15299, 15399-15499, 15599-15699, 15799-15899, 15999-16099, 16199-16299, 16399-16499, 16599-16699, 16799-16899, 16999-17099, 17199-17299, 17399-17499, 17599-17699, 17799-17899, 17999-18099, 18199-18299, 18399-18499, 18599-18699, 18799-18899, 18999-19099, 19199-19299, 19399-19499, 19599-19699, 19799-19899, 19999-20099, 20199-20299, 20399-20499, 20599-20699, 20799-20899, 20999-21099, 21199-21299, 21399-21499, 21599-21699, 21799-21899, 21999-22099, 22199-22299, 22399-22499, 22599-22699, 22799-22899, 22999-23099, 23199-23299, 23399-23499, 23599-23699, 23799-23899, 23999-24099, 24199-24299, 24399-24499, 24599-24699, 24799-24899, 24999-25099, 25199-25299, 25399-25499, 25599-25699, 25799-25899, 25999-26099, 26199-26299, 26399-26499, 26599-26699, 26799-26899, 26999-27099, 27199-27299, 27399-27499, 27599-27699, 27799-27899, 27999-28099, 28199-28299, 28399-28499, 28599-28699, 28799-28899, 28999-29099, 29199-29299, 29399-29499, 29599-29699, 29799-29899, 29999-30099, 30199-30299, 30399-30499, 30599-30699, 30799-30899, 30999-31099, 31199-31299, 31399-31499, 31599-31699, 31799-31899, 31999-32099, 32199-32299, 32399-32499, 32599-32699, 32799-32899, 32999-33099, 33199-33299, 33399-33499, 33599-33699, 33799-33899, 33999-34099, 34199-34299, 34399-34499, 34599-34699, 34799-34899, 34999-35099, 35199-35299, 35399-35499, 35599-35699, 35799-35899, 35999-36099, 36199-36299, 36399-36499, 36599-36699, 36799-36899, 36999-37099, 37199-37299, 37399-37499, 37599-37699, 37799-37899, 37999-38099, 38199-38299, 38399-38499, 38599-38699, 38799-38899, 38999-39099, 39199-39299, 39399-39499, 39599-39699, 39799-39899, 39999-40099, 40199-40299, 40399-40499, 40599-40699, 40799-40899, 40999-41099, 41199-41299, 41399-41499, 41599-41699, 41799-41899, 41999-42099, 42199-42299, 42399-42499, 42599-42699, 42799-42899, 42999-43099, 43199-43299, 43399-43499, 43599-43699, 43799-43899, 43999-44099, 44199-44299, 44399-44499, 44599-44699, 44799-44899, 44999-45099, 45199-45299, 45399-45499, 45599-45699, 45799-45899, 45999-46099, 46199-46299, 46399-46499, 46599-46699, 46799-46899, 46999-47099, 47199-47299, 47399-47499, 47599-47699, 47799-47899, 47999-48099, 48199-48299, 48399-48499, 48599-48699, 48799-48899, 48999-49099, 49199-49299, 49399-49499, 49599-49699, 49799-49899, 49999-50099, 50199-50299, 50399-50499, 50599-50699, 50799-50899, 50999-51099, 51199-51299, 51399-51499, 51599-51699, 51799-51899, 51999-52099, 52199-52299, 52399-52499, 52599-52699, 52799-52899, 52999-53099, 53199-53299, 53399-53499, 53599-53699, 53799-53899, 53999-54099, 54199-54299, 54399-54499, 54599-54699, 54799-54899, 54999-55099, 55199-55299, 55399-55499, 55599-55699, 55799-55899, 55999-56099, 56199-56299, 56399-56499, 56599-56699, 56799-56899, 56999-57099, 57199-57299, 57399-57499, 57599-57699, 57799-57899, 57999-58099, 58199-58299, 58399-58499, 58599-58699, 58799-58899, 58999-59099, 59199-59299, 59399-59499, 59599-59699, 59799-59899, 59999-60099, 60199-60299, 60399-60499, 60599-60699, 60799-60899, 60999-61099, 61199-61299, 61399-61499, 61599-61699, 61799-61899, 61999-62099, 62199-62299, 62399-62499, 62599-62699, 62799-62899, 62999-63099, 63199-63299, 63399-63499, 63599-63699, 63799-63899, 63999-64099, 64199-64299, 64399-64499, 64599-64699, 64799-64899, 64999

Search completed: July 18, 2001, 17:29:07
Job time: 99 sec



GenCore version 4.5
Copyright (c) 1994 - 2000 Compugen Ltd

CM protein - protein search, using sw model

Run on: July 18, 2001, 17:27:28 ; Search time 11.67 seconds
(without alignments)
375.724 Million bits of data, 300

Title: US-09-463-480-4

Perfect score: 685

Sequence: 1 MRVAVPFACVLCFQVHKAA.....KYAVDVPSPSTLSSDINS 128

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 94435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	74.5	10.9	1115	1	YAA4_SCHPO	Q07988 schizosacch
2	71	10.4	593	1	YF05_MYCTU	Q57333 mycobacteri
3	69	10.1	833	1	DL_DROME	P10041 Drosophila
4	69	10.1	1788	1	YF72_CABLL	Q07223 baccharid
5	68	9.9	187	1	VGG_BPAL3	P01281 bacterioph
6	68	9.9	860	1	LDLR_HUMAN	P01130 homo sapien
7	67	9.8	617	1	FXK1_MOUSE	P42128 mus musculu
8	67	9.8	648	1	TKT_MYCPN	P75611 mycoplasma
9	66.5	9.7	216	1	O5P4_HUMAN	Q06652 homo sapien
10	66	9.6	146	1	DER2_DERPT	P49278 dermatophag
11	66	9.6	276	1	PIPR_PASHA	Q08869 pasteurella
12	66	9.5	531	1	PEPF_ASPNG	P52718 aspergillus
13	66	9.5	662	1	DCHS_HUMAN	P19113 homo sapien
14	65	9.5	255	1	EL3B_SOVHR	P52295 elytrine max
15	65	9.5	381	1	NCK2_HUMAN	Q43639 homo sapien
16	65	9.5	858	1	ENV_HV2RO	P04577 human immun
17	65	9.5	1186	1	CEAA_BACTIS	Q45710 bacillus th
18	64.5	9.4	643	1	Y931_MELIA	Q53113 melibacter
19	64.5	9.4	738	1	YAS5_SCHPO	Q10145 schizosacch
20	64.5	9.4	863	1	LWPF_CHICK	P48165 gallus gall
21	64.5	9.4	1664	1	SLP1_CLODM	Q06952 clostridium
22	64	9.3	951	1	HEX_ADE05	P04133 human atoxo
23	64	9.3	1465	1	AT7B_HUMAN	P55670 homo sapien
24	63.5	9.3	407	1	AK2_BACST	P53553 bacillus st
25	63.5	9.3	701	1	YIA5_YEAST	P40557 saccharomye
26	63.5	9.3	970	1	T3RE_RPPI	P08764 bacterioph
27	63	9.2	573	1	H1FA_HAELN	P11729 haemophilu
28	63	9.2	346	1	POS_PAT	P17438 patiscent
29	62.5	9.1	321	1	AENA_ASFNC	P42256 aspergillus
30	62.5	9.1	327	1	OBAL_HUMAN	P45222 homo sapien
31	62.5	9.1	429	1	GRT1_CAPPL	P11776 capnocytoph
32	62.5	9.1	854	1	LDLR_CRIGR	P45950 cricetus
33	62	9.1	198	1	Y569_PRRHO	G58944 pyrodictus

RESULT 1

ID	YAA4_SCHPO	STANDARD	PRT	1115 AA
AC	Q07988			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-NOV-1996 (Rel. 34, Last annotation update)			
DE	HYPOTHETICAL 126.9 KDA PROTEIN C2257.04 IN CHROMOSOME I.			
GN	SPAC2257.04			
OS	Schizosaccharomyces pombe (Fission yeast).			
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;			
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;			
OC	Schizosaccharomycetes			
OX	NCBI_TaxID:4896			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-972			
RA	Redrock K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;			
RI	Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.			
CC	1- SIMILARITY: TO YEAST PAN2 AND TO C.ELEGANS F31E3.4.			

 CC this SWISS-prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.sib.ch/announc/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: 254328; CAA91128.1 ;
 DR InterPro: IPR000520 ;
 DR InterPro: IPR001394 ;
 DR Pfam: PF00929; Exonuclease; 1.
 DR Pfam: PF00443; DPH-2; 1.
 DR PROSITE: PS00225; DPH_2_3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1115 AA, 126919 MW, 26661518.972690 CRC64;

Query Match 10.9%; Score 74.5; DB 1; Length 1115;

Best Local Similarity 21.1%, Pred. No. 5.3;

Matches 32, Conservative 31, Mismatches 54, Indels 35, Gaps 8;

QY	2	PAAVAFACVLCFQVHKAAADRTCNPTDPMVTOTIG -LTIGSKQEEFNVLN----	55
Y4	102	ESMLLYSLT CEFIAVLISPSICINYPAMPYSPNNDFVWIDSYWQ-KITVVAAPG	159
QY	56	NIYAGSNVNSDGLHTEPT--DPHILRLPS-----DGTNN	91
Y4	160	ISHKRAKADVLTQEPH ELADDSISICILHFDNDFIVKRVAVITGLSIDLSCN	218
QY	92	CLVNRAPVSHAT LVAKRYAWVPTSTSI	121
Y4	219	ELLSQVYTHETVEMLVATVFWKCFRESSEIV	270

P15499 mus musculu
 Q84367 chlamydia t
 Q8743 bos taurus
 P23738 mus musculu
 P46103 plasmodium
 P53590 sus scrofa
 P18477 actinomyces
 P33235 escherichia
 G48042 bacterioph
 P45285 mus musculu
 P47312 mycoplasma
 P21657 saccharomye

RESULT 2

YP65_MYCTU STANDARD: PRG: 583 AA.
 AC Q50734;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DI 40-MAY-2000 (rel. 49, Last annotation update)
 DE HYPOHETICAL 6211 KDA PROTEIN RV2565;
 IN RV2565 OR MTC9904.06;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium; Mycobacterium.
 NCBI_TaxID=1774;
 RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN H37RV;
 RC MEDLINE 98295087; PubMed 944290;
 RA Cole S.T., Broese K., Parkhill J., Garnier T., Churcher C., Harris B.,
 RA Gordon S.V., Eigmeier K., Gars S., Barry C.E. III, Tekala E.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies P., DeLisle K., Feltwell T., Gentles S., Harris R., Hill D.,
 RA Horsby T., Jacobs K., Kadh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RI Nature 393:537-544(1998)
 CC -1- SIMILARITY: BELONGS TO THE UPF0028 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (see http://www.ebi.ac.uk/ncbi/ or
 CC or send an email to license@sib.ch)
 CC
 DR EMBL: Z77259; CAP01050.1;
 DR TuberColist: 892545;
 DR InterPro: IPR000595;
 DR InterPro: IPR001423;
 DR Pfam: PF01173; UPF0028; 1;
 DR Pfam: PF00027; CMBP BINDING_3; 1;
 DR PROSITE: PS50042; CMBP BINDING_3; 1;
 DR PROSITE: PS0127; UPF0028; 1;
 KW Hypothetical protein.
 SQ SEQUENCE 583 AA: 62123 MW: 65666062BAE762F7 CRC64;
 Query Match 10.4% Score 71; DB 1; Length 583;
 Best Local Similarity 29.0% Pred. No. 5.8;
 Matches 27; Conservative 13; Mismatches 41; Indels 22; Gaps 4;
 QY 52 NLTNNLYAGSRVSKVSCGLHTT... EPI DEHTIFLSDGTNRNLVRNAP 99
 11
 Db 462 NIPANWCAATCEVGVVIRRETVPSKPELIDIVIVVLLRLELLELQTRAL 517
 QY 100 ISNATVAHYVA WVVPSSTSTTSSIN 126
 11
 Db 518 IAPPTIRAPDIAASTANIPPTPRVAATTPPVS 550
 RESULT 3
 DL_PROMO STANDARD: PRG: 833 AA.
 AC P10041; Q9VDY2; 99108;
 DT 01-MAR-1989 (rel. 10, Created)
 DT 01-OCT-2000 (rel. 46, Last sequence update)
 DI 01-OCT-2000 (rel. 46, Last annotation update)
 DE NEUROGENIC LOCUS DELTA PROTEIN PRECURSOR.
 GN DL OR C53619.
 OS Drosophila melanogaster (Fruit fly).

OC Eukaryota, Metazoa, Arthropoda, Tracheata, Hexapoda, Insecta;
 OC Pterygota; Neoptera, Endopterygota, Diptera; Brachycera, Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RP TISSUE Embryo;
 RC MEDLINE 89196890; PubMed 3149249;
 RA Kozegowski C., Altan A.K., Pfeiffer K., Koch P., Muskaritch M.A.T.;
 RT "Delta, a Drosophila neurogenic gene, is transcriptionally complex and
 RT encodes a protein related to blood coagulation factors and epidermal
 RT growth factor of vertebrates."
 RI Science 251:1724-1735(1988).
 RN [2]
 SEQUENCE FROM N.A.
 RP STRAIN Oregon R; TISSUE-Embryo;
 RC MEDLINE 89196890; PubMed 3149249;
 RA Kozegowski C., Altan A.K., Pfeiffer K., Koch P., Muskaritch M.A.T.;
 RT "Delta, a Drosophila neurogenic gene, is transcriptionally complex and
 RT encodes a protein related to blood coagulation factors and epidermal
 RT growth factor of vertebrates."
 RI Science 251:1724-1735(1988).
 RN [3]
 SEQUENCE FROM N.A.
 RP STRAIN BEKKELEY;
 RC MEDLINE 20196006; PubMed 10741142;
 RA Adams M.D., Colniker S.E., Holt R.A., Evans C.A., Geyrhofer J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallo R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.C., Holt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Artzy-Baron A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballwey L.M., Basu A., Bazzucchi J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bonos P.V., Borham R.P., Boudreau N., Boudreau S.,
 RA Bracke J., Brackham M.E., Brack J., Brokstein P., Brothier P.,
 RA Bruford A.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Dayonport L.R., Davies P.,
 RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dong L.E., Downes M., Dudan Kcha S., Durkin R.C., Dunn P.,
 RA Durkin K.J., Ewanicki C.C., Ferrara C., Ferrera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garq N.S., Gilbert W.M., Glasser K.,
 RA Glöckle A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiland T.J., Hernandez J.R., Brock J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Iqbal M.,
 RA Zaidi M., Kishikawa F., Karpen G.H., Ko Z., Kemsion J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira G.D., Kraft G., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei R., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy P., Murphy L., Muzay D.M., Nelson D.L.,
 RA Nelson D.K., Nelson K.A., Nixon K., Nusskern D.R., Pachter J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollar J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Staden-Kiamos L., Simpson M., Skupski M.P., Smith T.,
 RA Spier F., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swierkas K., Tector P., Turner P., Venter E., Wang A.H., Wang X.,
 RA Wang Z., Wasserman P.A., Weinstock G.M., Weissmanbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye Y., Yeh F., Zeng J., Zhang M., Zhang X., Zhang Y., Zhou L.,
 RA Zhou X.H., Zhou F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RI Science 287:2185-2195(2000).
 RN [4]
 SEQUENCE OF 422-621 FROM N.A.
 RP TISSUE Embryo;
 RC MEDLINE 87218547; PubMed 3107986;
 RA Knust E., Dietrich U., Tepass U., Bremer K.A., Weigold D., Vaesslin H.,
 RA Campos-ortega J.A.;
 RT "EGF homology sequences encoded in the genome of Drosophila
 RT melanogaster, and their relation to neurogenic genes."
 RI EMBO J. 6:761-766(1987).
 RN [5]

OS Caenorhabditis elegans.
 OC Eukaryotes; Metazoa; Rotatoria; Chromadorea; Platyhelminthes.
 CC Khabditidae; Polidromina; Caenorhabditis.
 CX NCBI_TaxID: 6230.
 FN [1]
 RP
 RA STRAIN: RW1501, N2.
 RL Submitted (MAR-1995) to the EMBL/GenBank/Trna databases.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch>)
 CC or send an email to license@isb-sib.ch.
 CC
 DR EMBL: D23169; AAC08800.1.
 DR WormPop: B0228.2; CE0744.
 KW Hypothetical protein.
 SQ SEQUENCE 1788 AA; 200633 MW; 8F04D4BAE19B20A2 CRC64;
 Query Match 10.1%; Score 69; DB 1; Length 1788;
 Best Local Similarity 28.9%; Pred. No. 45;
 Matches 24; Conservative 10; Mismatches 27; Indels 22; Gaps 4;
 QY 18 KALADKTKN-----PTDMVQI-----TLLTGGKGVFVNLINNYIAQS 62
 DB 1006 RAGFADEKFADEFAKIGKDAQFSAIPIQNEIKFNEICASGEISVIEETVCEKFS 1065
 QY 63 NRVKVSIGLHTI---EPIDRHII 82
 DB 1066 ---SFEETSITKVAKNIEPHVI 1084
 RESULT 5
 VQ6_BVAL3
 ID VQ6_BVAL3 STANDARD; PRT; 187 AA.
 AC P41281;
 DT 01-JUL-1994 (Rel. 26, Created)
 DI 01-JUL-1994 (Rel. 26, Last sequence update)
 DE 15-JUL-1994 (Rel. 36, Last annotation update)
 DE MAJOR SPIKE PROTEIN (G PROTEIN) (GPC).
 GN G.
 OS Bacteriophage alpha-4.
 OC Viruses; ssDNA viruses; Microviridae; Microvirus.
 CX NCBI_TaxID: 10849;
 FN [1]
 RP SEQUENCE FROM N A
 RX MEDLINE:92224109; PubMed:1532998.
 RA Kodaira K.-I., Nakano K., Okada S., Taketo A.;
 RT "Nucleotide sequence of the genome of the bacteriophage alpha 4;
 RT interrelationship of the genome structure and the gene products with
 RT those of the phages, phi X174, G4 and phi K8";
 RL Biochim. Biophys. Acta 1133:277-288(1992).
 RN [2]
 RP PARTIAL SEQUENCE.
 RX MEDLINE:80049950; PubMed:387790;
 RA Sims J., Capon D., Dressler L.;
 RT "dnaG (primase)-dependent origins of DNA replication. Nucleotide
 RT sequences of the negative strand initiation sites of bacteriophages
 RT SF-1, phi K, and alpha 4";
 RL J. Biol. Chem. 254:12615-12628(1979).
 CC -1- FUNCTION: MAJOR SPIKE COMPONENT. INVOLVED IN THE ATTACHMENT TO
 CC THE BACTERIAL HOST.
 CC -2- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES OF THE G PROTEIN, AND
 CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
 CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch>)
 CC or send an email to license@isb-sib.ch.
 CC
 DR EMBL: X63322; CAA12882.1;
 DR PIR: C04253; C04253.
 DR PIR: S22331; S22331.
 DR BSSP: P04643; IAL0.
 KW Coat protein.
 SQ SEQUENCE 187 AA; 19565 MW; 6604DB964F7758EF CRC64;
 Query Match 9.5%; Score 68; DB 1; Length 187;
 Best Local Similarity 28.0%; Pred. No. 32;
 Matches 28; Conservative 11; Mismatches 49; Indels 22; Gaps 5;
 QY 47 LLLGLLGGKQQLIVRLIN RLYCAASVSKVSKTGLH -----TPEITDP 79
 DB 19 SVTGVIPAAPFQTHIVNGSIIAERAVNNLY AMNVSTSSAGSIVAMKVDTSEIDP 77
 QY 80 HILPLSLDGINNLYNNIAPISDAIIIAFKVAMDVPESES 119
 DB 78 NCV--ISAGVNLSEFAGTSYPI--VGLVRFESASEQPIA 114
 RESULT 6
 LDLR_HUMAN
 ID LDLR_HUMAN STANDARD; PRT; 860 AA.
 AC P01130;
 DT 21-JUL-1986 (Rel. 01, Created)
 DI 21-JUL-1986 (Rel. 01, Last sequence update)
 DE 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LOW-DENSITY LIPOPROTEIN RECEPTOR PRECURSOR (LDL RECEPTOR).
 GN LDLR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID: 9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:85024898; PubMed:6091915;
 RA Yamamoto T., Davis C.G., Brown M.S., Schneider W.L., Casey M.L.,
 RA Goldstein J.L., Russell D.W.;
 RT "The human LDL receptor: a cysteine rich protein with multiple Ala
 RT sequences in its mRNA";
 RL Cell 39:27-38(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 85218750; PubMed 2988123;
 RA Suckhot T.G., Goldstein J.L., Brown M.S., Russell D.W.;
 RT "The LDL receptor gene: a mosaic of exons shared with different
 RT proteins";
 RL Science 228:815-822(1985).
 RN [3]
 RP NCTADINESIS OF CYTOPLASMIC DOMAIN.
 RX MEDLINE:87166314; PubMed:3104386;
 RA Davis C.G., van Griel J.F., Russell D.W., Brown M.S., Goldstein J.L.;
 RT "The low density lipoprotein receptor: identification of amino acids
 RT in cytoplasmic domain required for rapid endocytosis";
 RL J. Biol. Chem. 262:4075-4082(1987).
 RN [4]
 RP O-GLYCOSYLATION DOMAIN.
 RX MEDLINE 96140336; PubMed:3305267;
 RA Davis C.G., Ellhammer A., Russell D.W., Schneider W.L., Kornfeld S.;
 RT "Detection of clustered O-linked carbohydrates does not impair
 RT function of low density lipoprotein receptor in transfected
 RT fibroblasts";
 RL J. Biol. Chem. 261:2828-2838(1986).
 RN [5]
 RP STRUCTURE BY NMR OF 20-67.

FX MEDLINE=95327641; PubMed=7603001;
 FA Daly N.F., Scallan M.J., Djordjevic J.T., Kroon P.A., Smith P.;
 FT "Three-dimensional structure of a cysteine-rich repeat from the low-
 density lipoprotein receptor.";
 FL Proc. Natl. Acad. Sci. U.S.A. 92:6334-6338(1995)
 FN [6]
 RP STRUCTURE BY NMP OF 65-104
 FX MEDLINE=96062511; PubMed=7578052;
 FA Daly N.L., Djordjevic J.T., Kroon P.A., Smith P.;
 FT "Three-dimensional structure of the second cysteine-rich repeat from
 the human low-density lipoprotein receptor";
 FL Biochemistry 34:14474-14481(1995)
 FN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 136-242
 FX MEDLINE=97404319; PubMed=9262405;
 FA Fass D., Blacklow S.C., Kim P.S., Beigel Y.M.;
 FT "Molecular basis of familial hypercholesterolemia from structure of
 LDL receptor module";
 FL Nature 388:691-693(1997)
 FN [8]
 RP REVIEW ON FH VARIANTS.
 FX MEDLINE=93750847; PubMed=1301956;
 FA Hobbs H.H., Brown M.S., Goldstein J.L.;
 FT "Molecular genetics of the LDL receptor gene in familial
 hypercholesterolemia.";
 FL Hum. Mutat. 1:445-466(1992)
 FN [9]
 RP REVIEW ON FH VARIANTS.
 FX MEDLINE=97569482; PubMed=9216531;
 FA Varret M., Fares J.-P., Colod-Beroud G., Junien J., Bellet D.;
 FT "Software and database for the analysis of mutations in the human LDL
 receptor gene";
 FL Nucleic Acids Res. 25:172-180(1997)
 FN [10]
 RP VARIANTS FH AFRIKANER-1; -2 AND -3.
 FX MEDLINE=89340944; PubMed=2569482;
 FA Leitersdorf E., van der Westhuyzen D.P., Coetzee G.A., Biers H.H.;
 FT "Two common low density lipoprotein receptor gene mutations cause
 familial hypercholesterolemia in Afrikaners.";
 FL J. Clin. Invest. 84:954-961(1989)
 FN [11]
 RP VARIANT FH BARI.
 FX MEDLINE=86161691; PubMed=3955557;
 FA Davis C.G., Lehrman M.A., Russell D.W., Anderson R.G.W., Brown M.S.,
 FA Goldstein J.L.;
 FT "The J.E. mutation in familial hypercholesterolemia: amino acid
 substitution in cytoplasmic domain impedes internalization of LDL
 receptors.";
 FL Cell 45:15-24(1986)
 FN [12]
 RP VARIANTS FH DIPPAN-1 AND -2.
 FX MEDLINE=93450005; PubMed=8347889;
 FA Rubinstein D.C., Tajima I., Leitersdorf E., Coetzee G.A.,
 FA van der Westhuyzen D.P.;
 FT "Identification of two new LDL receptor mutations causing homozygous
 familial hypercholesterolemia in a South African of Indian origin";
 FL Biochim Biophys. Acta 1192:75-82(1993)
 FN [13]
 RP VARIANTS FH FRENCH CANADIAN-2; -3 AND -4.
 FX MEDLINE=90203205; PubMed=2318661;
 FA Leitersdorf E., Robin F.J., Davignon J., Hobbs H.H.;
 FT "Common low-density lipoprotein receptor mutations in the French
 Canadian population";
 FL J. Clin. Invest. 85:1014-1024(1990)
 FN [14]
 RP VARIANT FH OSAKA-3.
 FX MEDLINE=93078772; PubMed=1446662;
 FA Miyake Y., Tajima S., Funahashi T., Yamamura T.,
 FT "A point mutation of low-density-lipoprotein receptor causing rapid
 degradation of the receptor.";
 FL Eur. J. Biochem. 210:17(1992)
 FN [15]
 RP VARIANTS FH SWEDISH LYS-277, THR-423 AND ASN-579.
 FX MEDLINE=95362239; PubMed=7635461;
 FN [16]
 RP VARIANT FH PISCATAWAY.
 FX MEDLINE=91428152; PubMed=1867500;
 FA Meiner V., Landsberger D., Berkman N., Peshef A., Sedal P.,
 FA Seftel H.C., van der Westhuyzen D.P., Joensuu M.S., Coetzee G.A.,
 FA Leitersdorf E.;
 FT "A common Lithuanian mutation causing familial hypercholesterolemia
 in Ashkenazi Jews";
 FL Am. J. Hum. Genet. 49:443-449(1991)
 FN [17]
 RP VARIANT FH SAFED
 FX MEDLINE=93216279; PubMed=8462973;
 FA Leitersdorf E., Peshef A., Meiner V., Dann E.J., Beigel Y.,
 FA van Poggen F.G., van der Westhuyzen D.P., Coetzee G.A.;
 FT "A missense mutation in the low density lipoprotein receptor gene
 causes familial hypercholesterolemia in Sephardic Jews.";
 FL Hum. Genet. 91:141-147(1993)
 FN [18]
 RP VARIANT FH TRIESTE.
 FX MEDLINE=94224719; PubMed=8168830;
 FA Felli N., Garuti P., Pedrazzi P., Ghisellini M., Simone M.I.,
 FA Tiozzo R., Cattin L., Valenti M., Rrolleri M., Bertolini S.,
 FA Stefanutti C., Calandra S.;
 FT "A new missense mutation (Cys297>>Phe) of the low density
 lipoprotein receptor in Italian patients with familial
 hypercholesterolemia (FHTRieste).";
 FL Hum. Genet. 93:538-540(1994)
 FN [19]
 RP VARIANT FH ZAMBIA.
 FX MEDLINE=89264573; PubMed=2705768;
 FA Soutar A.K., Knight R.L., Patel D.D.;
 FT "Identification of a point mutation in growth factor repeat C of the
 low density lipoprotein receptor gene in a patient with homozygous
 familial hypercholesterolemia that affects ligand binding and
 intracellular movement of receptors.";
 FL Proc. Natl. Acad. Sci. U.S.A. 86:4166-4170(1989)
 FN [20]
 RP VARIANTS FH PORT HIS-401 AND TURK ASP-844
 FX MEDLINE=96029270; PubMed=7573037;
 FA Kaivisto U.-M., Viikari J.S., Kontula K.;
 FT "Molecular characterization of minor gene rearrangements in Finnish
 patients with heterozygous familial hypercholesterolemia:
 identification of two common missense mutations (Gly823>>Asp and
 Leu380>>His) and eight rare mutations of the LDL receptor gene";
 FL Am. J. Hum. Genet. 57:789-797(1995)
 FN [21]
 RP VARIANTS FH LYS-140; SER-328 AND LEU-685.
 FX MEDLINE=9601600; PubMed=7883548;
 FA Maruyama T., Miyake Y., Tajima S., Harada-Shiba M., Yamamura T.,
 FA Tsushima M., Kishino K.-I., Higuchi Y., Funahashi T., Matsuzawa Y.,
 FA Yamamoto A.;
 FT "Common mutations in the low-density-lipoprotein-receptor gene
 causing familial hypercholesterolemia in the Japanese population.";
 FL Arterioscler Thromb. Vasc. Biol. 15:1713-1718(1995)
 FN [22]
 RP VARIANT FH FRENCH HIS-564.
 FX MEDLINE=96055524; PubMed=7550239;
 FA Lricor-Guier F., Sainet-Jore H., Valenti K., Foulon T., Bost M.,
 RA Hadjian A.J.;
 FT "Identification of a mutation, N543H, in exon 11 of the low-density
 lipoprotein receptor gene in a French family with familial
 hypercholesterolemia";
 FL Hum. Mutat. 6:87-88(1995)
 FN [23]
 RP VARIANTS FH SWEDISH LYS-277, THR-423 AND ASN-579.
 FX MEDLINE=95362239; PubMed=7635461;
 FN [24]


```

DR LMBL: AF276249; AAF86462.1; -
DR PDB: 1A9V; 14-OCT-98.
KW Allergen; Signal; 3D-structure; Polymorphism.
FT SIGNAL 1 17
FT CHAIN 18 146 MITE ALLERGEN DER P 2.
FT DISULFID 25 136
FT DISULFID 38 44
FT DISULFID 50 95
FT VARIANT 43 39
FT VARIANT 40 40
FT VARIANT 44 44
FT VARIANT 47 47
FT VARIANT 49 49
FT VARIANT 56 56
FT VARIANT 61 61
FT VARIANT 75 75
FT VARIANT 78 78
FT VARIANT 81 81
FT VARIANT 95 95
FT VARIANT 98 98
FT VARIANT 108 108
FT VARIANT 111 111
FT VARIANT 114 114
FT VARIANT 116 116
FT VARIANT 118 118
FT VARIANT 127 127
SQ SEQUENCE 146 AA; 15999 MW; 59182FA7F9D663AF C9C64;

Query Match 9.6%; Score 66; DB 1; Length 146;
Best Local Similarity 29.4%; Pred. No. 3.9;
Matches 20; Conservative 6; Mismatches 28; Indels 14; Gaps 3;

QY 49 FEVNLINLLYCAQSNVYKSCIRGLHTTEP-IDPHILRLSDGNTNCLVNGAPISIALIVA 107
DB 58 FRAN--QNTKTAKIETKASITGLLEVLVPEJIDPNACHTMK-----GPIVKGGQGYD 104
QY 108 FRYAMVP 115
DB 105 IKYTNPV 112

RESULT 11
PLPB_PASHA
ID PLPB_PASHA STANDARD; PRT: 276 AA.
AC Q08869; Q07364;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE OUTER MEMBRANE LIPOPROTEIN 2 PRECURSOR (PLP2).
GN PLPB.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
OC Mannheimia.
OX NCBI_TaxID:75985;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN SECTYPE A1;
RX MEDLINE:93328110; PubMed:8335249;
RA Murphy G.L., Whitworth L.C.;
RT "Analysis of tandem, multiple genes encoding 30 kDa membrane proteins
in Pasteurella haemolytica A1."
RL Gene 129:107-111(1993).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN SECTYPE A1;
RX MEDLINE:94011478; PubMed:8408866;
RA Gooney B.J., Lo R.Y.C.;
RT "Three contiguous lipoprotein genes in Pasteurella haemolytica A1
which are homologous to a lipoprotein gene in Haemophilus influenzae
type b."
RL Infect. Immun. 61:4682-4688(1993).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID

ANCHOR (PROBABLE).
-!- SIMILARITY: BELONGS TO THE NLPA FAMILY OF LIPOPROTEINS.
This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, usage by and for commercial
entities requires a license agreement (See: http://www.ebi.ac.uk/ebis/ebis/
or send an email to license@ebi.ac.uk).
EMBL: L11037; AAA25539.1; -
DR EMBL: L11627; AAA25547.1; -
DP PROSITE: PS00013; PROKAP_LipoPROTEIN; 1
KW Outer membrane; Lipoprotein; Signal; Multigene family.
FT SIGNAL 1 19
FT CHAIN 20 276
FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
FT CONFLICT 268 270 FNG -> LTVH (IN PEF 2).
SQ SEQUENCE 276 AA; 30157 MW; DF8F4434ADBD4950 CRC64;

Query Match 9.6%; Score 66; DB 1; Length 276;
Best Local Similarity 25.8%; Pred. No. 8.1;
Matches 16; Conservative 11; Mismatches 45; Indels 0; Gaps 0;

QY 20 ALAPKNTPTFTMTQTITRTTICGKCEPTVNLINLYCAQSNVYKSCIRGLHTTEPDP 79
DB 172 ATENDLIENPKNIKIVQADTSILTEMLDDVGLAVINNTYCAQLSDKELIVESKDSF 231
QY 80 HI 81
DB 232 YV 233

RESULT 12
PEPF_ASPEG
ID PEPF_ASPEG STANDARD; PRT: 531 AA.
AC P52718;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE SERINE-TYPE CARBOXYPEPTIDASE 2 PRECURSOR (EC 3.4.16.-) (PROTEINASE F)
DE (CPD-11).
GN PEPF.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eriotheliales;
OC Eriotheliales; Trichocomaceae; Mitosporic Trichocomaceae; Aspergillales.
OX NCBI_TaxID:5061;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN-GPS 120.49 / N400;
RX MEDLINE:95129695; PubMed 7828508;
RA van den Hembergh J.P.T.W., Jaraí G., Ruxton F.P., Visser J.;
RT "Cloning, characterization and expression of pepF, a gene encoding a
serine carboxypeptidase from Aspergillus niger."
RL Gene 151:73-75(1994).
RN 121
RP REVISIONS.
RA Schaap P.J., Visser J.;
RL Submitted (OCT 1997) to the EMBL/GenBank/DBJ databases.
RN 131
RP SEQUENCE OF 53 71, AND CHARACTERIZATION.
RX MEDLINE:92344370; PubMed:1637154;
RA Dal Began P., Ribadeau-Dumas R., Broddam K.;
RT "Purification and characterization of two serine carboxypeptidases
from Aspergillus niger and their use in C-terminal sequencing of
proteins and peptide synthesis."
EL Appl. Environ. Microbiol. 59:2114-2152(1992).
CC -!- FUNCTION: REMOVES ANY AMINO ACID FROM THE C-TERMINUS OF A LONG
PEPTIDE. DIGESTS PREFERENTIALLY PEPTIDES CONTAINING A POSITIVELY
CHARGED RESIDUE IN P1' POSITION, AS WELL AS ARGININE, LYSINE OR
PHENYLALANINE IN P1 POSITION OF ESTER SUBSTRATE. OPTIMAL PH IS 4;

```

CC ENZYME IS UNSTABLE ABOVE pH 8. CATALYZES ALCO PEPTIDE SYNTHESIS.
 CC -|- CATALYTIC ACTIVITY: RELEASE OF A C-TERMINAL AMINO ACID WITH A
 CC BROAD SPECIFICITY.
 CC -|- ENZYME REGULATION: INHIBITED BY LFP, ANH, HG(CL)2.
 CC -|- SUBUNIT: MONOMER.
 CC -|- INDUCTION: IN THE FOLLOWING GROWTH CONDITIONS: ACIDIC PH, ABSENCE
 CC OF NITROGEN OR CARBON SOURCE
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 510. ALFA PR 40 47 70
 CC SERINE CARBOXYPEPTIDASE FAMILY
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announcements/>
 CC or send an email to license@isb.ch).
 CC
 CC EMBL: X79541; CAA56075.1; -;
 CC EMBL: L33408; AAB57723.1; -;
 CC HSSP: P10619; 11VV;
 CC MEROPS: S10.005; -;
 CC InterPro: IPR001563; -;
 CC Pfam: PF00450; serine_carpt; 2;
 CC PRINTS: PR00724; CARBOXYPTASEC;
 CC PROSITE: PS00131; CARBOXYPEPT_SER_SF; 1;
 CC PROSITE: PS00560; CARBOXYPEPT_SER_HIS; 1;
 CC K W Hydrolyase; Carboxypeptidase; Glycoprotein, Cysteine, Signal.
 CC FT SIGNAL 1 25 POTENTIAL.
 CC FT PROPEP 26 52
 CC FT CHAIN 53 531 SERINE-TYPE CARBOXYPEPTIDASE P.
 CC FT ACT_SITE 211 211 BY SIMILARITY.
 CC FT ACT_SITE 430 430 BY SIMILARITY.
 CC FT ACT_SITE 507 507 BY SIMILARITY.
 CC FT CARBOHYD 20 20 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 63 63 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 94 94 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 155 155 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 228 228 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 271 271 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 309 309 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 378 378 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 436 436 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CARBOHYD 444 444 N-LINKED (GLCNAC...) (POTENTIAL).
 CC FT CONFLICT 53 53 I -> A (IN REF. 3).
 CC FT CONFLICT 57 57 D -> F (IN REF. 3).
 CC FT CONFLICT 63 63 N -> Y (IN REF. 3).
 CC FT CONFLICT 68 68 Y -> D (IN REF. 3).
 CC SQ SEQUENCE 531 AA; 59536 MW; 60DF0FA9E8478550 CRC64.

Query Match 9.6%; Score 66; DR 1; Length 531.
 Best Local Similarity 26.6%; Pred. No. 18;
 Matches 21; Conservative 16; Mismatches 24; Indels 18; Gaps 4;

QY 59 CAGSNVKKVSCD-----GLHTETFTDHPHLEFSSSTNRLVNNR3 418GAILVAF L-4
 DQ 384 CSVESVFVSGGCGPQEGEYSANPIF-HVLPQVTEINPVLVINGDYDMVLTINCTLLST 442
 QY 109 -----KYAWDVPISFSI 120
 DQ 443 QNMTWCKLGFDTAPSTPI 461

RESULT 13
 DCHS_HUMAN STANAPR; PRO; 552 AA.
 AC P19113;
 ET 01-NOV-1990 (Rel. 16, created)
 ET 01-FEB-1996 (Rel. 33, last sequence update)
 ET 01-OCT-2000 (Rel. 40, last annotation update)
 DE HISTIDINE DECARBOXYLASE (EC 4.1.1.22) (HDC).
 GN HDC.

QY Homo sapiens (Human).
 AC Eukaryota; Metazoa; Chordata; Claciata; Vertebrata; Euteleostomi;
 AC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 QX KBTJtaxid:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 94117471; PubMed 2216786;
 FA Yamauchi K., Saitoh K., Ohtsuka Y., Takeda Y., MacGowan K.,
 RA Watanabe T., Saitoh K., Yoshizawa M., Shibahara S., Takishima T.,
 RT "Nucleotide sequence of the cDNA encoding L-histidine decarboxylase
 derived from human basophilic leukemia cell line, KO-812-F.1".
 RC Nucleic Acids Res. 18:5891-5891(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 92119428; PubMed 1768863;
 FA Ohtsuka Y., Saitoh K., MacGowan K., Joseph L.F.;
 RT "Cloning of the cDNA encoding human histidine decarboxylase from an
 RT erythroleukemia cell line and mapping of the gene locus to chromosome
 RC 15".
 RC DNA Seq. 1:395-400(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 94117478; PubMed 8289422;
 FA Yatsunami K., Ohtsuka Y., Itochikawa M., Hiyoshi T., Ishibashi K.,
 RA Shida A., Shima Y., Nakagawa S., Yamauchi K., Yamamoto M.,
 RT Hayashi N., Watanabe T., Ichikawa A.;
 RC "Structure of the L-histidine decarboxylase gene".
 RC J. Biol. Chem. 269:1554-1559(1994).
 CC -|- CATALYTIC ACTIVITY: L-HISTIDINE - HISTAMINE + CO(2).
 CC -|- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -|- PATHWAY: BIOSYNTHESIS OF THE NEUROTRANSMITTER HISTAMINE.
 CC -|- SUBUNIT: HOMODIMER.
 CC -|- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (HDC, GAD, HDC AND
 CC TYROC).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announcements/>
 CC or send an email to license@isb.ch).
 CC
 CC EMBL: X54297; CAA38196.1; -;
 CC EMBL: M60445; AAC1698.1; -;
 CC EMBL: D16583; BAA04015.1; -;
 CC PIR: S11492; S11492.
 CC MIM: 142704; -;
 CC InterPro: IPR002129; -;
 CC Pfam: PF00282; pyridoxal_dec; 1.
 CC PRINTS: PR00800; YHDCRBOXYLASE.
 CC PROSITE: PS00392; HDC_GAD_HDC_YDC; 1.
 CC KW Lyase, Decarboxylase, Catecholamine biosynthesis, Pyridoxal phosphate.
 CC FT BINDING 305 305 PYRIDOXAL PHOSPHATE (POTENTIAL).
 CC FT CONFLICT 148 148 S -> Q (IN REF. 1).
 CC SQ SEQUENCE 352 AA; 74145 MW; 1751CFAA060P450 CRC64;

Query Match 9.6%; Score 66; DR 1; Length 662;
 Best Local Similarity 27.7%; Pred. No. 23;
 Matches 31; Conservative 16; Mismatches 45; Indels 20; Gaps 7;

QY 22 ADERTNPPTDENVTO-TTGTTCGRQKQEFVNL-INNLVYAGSNVKKVSCGLHTETFDIP 79
 DQ 296 ADSPFTIRSKMMVHVDPTGTWVKDKYKQQTESVNPVYLRHANSVATDMMHMDPLSR 355
 QY 80 H - - - - - LLEFLDGLNG LKVRREZALSHALLVAI FYAW-W 120
 DQ 355 PFSVPELKEVLEP-SFVFNL- - - - - ARVPPESTEMA-PYFESLVPNLSPEL 400

RESULT 14
 E136_SOVEN

```

ID 1414R SOYBN STANDARD: PRI: 255 AA.
AC P52195;
DI 01-OCT-1996 (rel. 44, Last sequence update)
DI 01-OCT-1996 (rel. 44, Last sequence update)
DI 01 NOV 1997 (rel. 45, Last annotation update)
DE GLUCAN ENDO-1,4-BETA-GLUCOSYLASE (EC 3.2.1.40) (C. G. BELA-SILVIAN
DE ENDOHYDROLASE) ((1-3)-BETA-GLUCANASE) (BETA 1,4-ENHANCERANASE)
DE (FRAGMENT).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucosids 1;
OC Fabales; Fabaceae; Papilionoideae; Glycine.
OX NCBI_TaxID 3847;
RN [1]
RP SEQUENCE FROM N.A.
RE STRAIN CV. CENTURY 84; TISSUE Leaf;
RX MEDLINE:95218610; PubMed:776605;
RA Kalinski A., Kowley D.L., Low D.S., Foley C., Rota G., Herman E.M.;
RT "Binding-protein expression is subject to temporal, developmental and
RI stress induced regulation in terminally differentiated soybean
FI nodules."
PL Planta 195;611-621(1995).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,3-BETA-D-GLUCOSIDIC LINKAGES
CC IN 1,3-BETA-D-GLUCANS.
CC -1- SIMILARITY: BELONGS TO FAMILY 17 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.scrib.ch)
CC or send an email to license@isb.scrib.ch.
CC
CC EMBL: U08415; AAA1955.1;
DR HSSP: P15747; IGHS.
DR InterPro: IPR000490;
DR Pfam: PF00332; Glyco_Hydro_17; 1.
DR PROSITE: PS00587; GLYCOSYL_HYDROL_F17; 1.
KW Hydrolase; Glycosidase; Multicene family.
FT NON_TER 1
FT ACT_SITE 183 184 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 255 AA: 27671 MW: 124248PTT183611C CRC64;

Query Match 9.5% Score 65; DR 1; Length 255;
Best Local Similarity 26.48; Pred. No. 9.5;
Matches 24; Conservative 19; Mismatches 25; Indels 20; Gaps 5;

QY 52 NLINLYCAQSNVKSVC-----DKLHTPEPLDPIHRLSDCTNRNINNG 97
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 65 NTONA1SSKKE1IKVSLAIDSLTILNSYDNGVETSD AEFY LKFI1 NPLVSNQ 118
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 98 AP1SHATLVAKRYKAD--VPPSESLISSDIN 126
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 119 APLANVYPYFAYANDQSIPLAYALTQOQN 149
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
NCK2 HUMAN
ID NCK2_HUMAN STANDARD: PRI: 381 AA.
AC G43639;
DI 01-OCT-2000 (rel. 40, Created)
DI 01-OCT-2000 (rel. 40, Last sequence update)
DI 01-OCT-2000 (rel. 40, Last annotation update)
DE CYTOSOLASMIC PROTEIN NCK2 (NCK ALAPTRK180FEIN 2) (CD253B3 ALAPTRK
DE PROTEIN NCK-BETA).
GN NCK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE:96411415; PubMed:9737977;
RA Chen M., She H., Davis E.M., Spicer C.M., Kim L., Ren K., Lebeau M.M.,
RA Li W.;
RT "Identification of NCK family genes, chromosomal localization,
RT expression, and signaling specificity."
RL J. Biol. Chem. 273;25191-25198(1998).
CC -1- FUNCTION: ADAPTOR PROTEIN WHICH ASSOCIATES WITH LYSG-LINE
CC PHOSPHORYLATED GROWTH FACTOR RECEPTORS OR THEIR CELLULAR
CC SUBSTRATES.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: URICITICUS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 SH3 DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.scrib.ch)
CC or send an email to license@isb.scrib.ch.
CC
CC EMBL: AF043119; AAC04841.1;
DR HSSP: P19174; IHSQ.
DR InterPro: IPR000980;
DR InterPro: IPR001452;
DR Pfam: PF00017; SH2_1;
DR Pfam: PF00018; SH3_3;
DR PROSITE: PS00452; SH3_DOMAIN;
DR PROSITE: PS00001; SH2_1;
DR PROSITE: PS00002; SH3_3;
KW SH3 domain; SH2 domain.
FT DOMAIN 2 62 SH3.
FT DOMAIN 121 171 SH3.
FT DOMAIN 196 258 SH3.
FT DOMAIN 286 380 SH2.
SQ SEQUENCE 481 AA: 44014 MW: 990588EA5AFA5714 CRC64;

Query Match 9.5% Score 65; DR 1; Length 481;
Best Local Similarity 29.6% Pred. No. 15;
Matches 18; Conservative 11; Mismatches 20; Indels 12; Gaps 3;

QY 21 LADDEKNTPTFMVTPFHTGTLTQGR-QEPEVNLINLYCAQSNVKSVCPLHTTEPLTP 79
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 410 LIPSESPSPFSVS-----LEASGPKKRFKVLGVDMVYV-----IQPRFTHMELVE 458
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 80 H 80
|
DB 359 H 359

Search completed: July 18, 2001, 17:29:24
Job time: 116 sec

```

Genome version 4.5

Copyright (c) 1993-2000 Computer Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2001, 17:27:28 : Search time 211 seconds
(without alignments)

802,609 million cell updates/sec

Title: US-09-463-480 4

Perfect score:

Sequences:

1 MRVAVFACVLCFQVHKA.....KFAVHICGILASDRES 123

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched:

425026 seqs, 13230527 residues

Total number of hits satisfying chosen parameters:

425026

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL_16:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_minc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_unclassified:*

13: sp_vertebrate:*

14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	685	100.0	128	10	Q91W7	Q91W7 oryza sativa
2	134	19.5	124	10	Q49384	Q49384 arabidopsis
3	129	18.9	123	10	Q49382	Q49382 arabidopsis
4	124	18.1	263	10	Q49381	Q49381 arabidopsis
5	116.5	17.0	1193	10	Q9MA43	Q9MA43 arabidopsis
6	78.5	11.5	419	11	Q70572	Q70572 mus musculus
7	77.5	11.3	835	11	Q92303	Q92303 mus musculus
8	77.5	11.3	835	11	Q90K6	Q90K6 mus musculus
9	76.5	11.2	333	10	Q9H30	Q9H30 arabidopsis
10	76	11.1	423	4	Q60906	Q60906 homo sapien
11	75.5	11.0	835	11	Q9QZ45	Q9QZ45 mus musculus
12	75	10.9	922	3	Q42861	Q42861 schistosom
13	74	10.8	1086	5	Q47302	Q47302 plasmodium
14	73	10.7	2349	2	P94750	P94750 escherichia
15	73	10.7	2383	2	P76347	P76347 escherichia
16	72.5	10.6	166	3	Q13522	Q13522 saccharomy
17	72	10.5	1261	10	Q9LYH5	Q9LYH5 arabidopsis
18	71	10.4	535	2	Q9X4P1	Q9X4P1 actinomyces
19	71	10.4	2842	4	Q10455	Q10455 homo sapien

ALIGNMENTS

RESULT 1

Q91W7

ID Q92PJ7

AC Q92PJ7

DT 01-MAY-1999

DT 01-MAY-1999

DT 01-MAY-1999

DE 1991

GN LIL1

OS Lillium longiflorum

SC Lillium longiflorum

OX NPRTaxID 4690

RN L11

RF SEQUENCE FROM N.A.

FX MED 1999:44762644

RA Xu H., Swoboda L., Bualla P.L., Singh M.B.

RT "Male gametic cell-specific gene expression in flowering plants."

RL Proc. Natl. Acad. Sci. U.S.A. 96:2554-2558(1999).

DR EMBL: AF110779; AAD19621; ...

SQ SEQUENCE 128 AA, 13851 MW, 18643 pI, 10.2 pI, 10.2 pI

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7

Q91W7


```

DR  PUTOPIA; LPR000483;
DR  InterPro; IPR000887;
DR  Info; PIR001611;
DR  Pfam; PF00560; LRR; 7;
DR  Pfam; PF01463; LRR; 1;
DR  Pfam; PF01582; TIR; 1;
DR  PRINTS; PR00019; LEUKICHRPT;
DR  PROSITE; PS00159; ALDOLASE_KHG_1; UNKNOWN_1;
KW  Receptor.
FI  VARIANT. 712 712 H -- P.
SQ  SEQUENCE 835 AA: 95558 MW: 969545911A4A0C17 CRC64;

Query Match
Best Local Similarity 24.94; Score 77.5; DB 11; Length 845;
Matches 39; Conservative 14; Mismatches 43; Indels 49; Gaps 8;

QY 21 LAUGKTCNTDPMVTGTTTGLTGKGFVEVNLNNLYAGSNVNVSG-----HGLH---- 72
DB 205 LSLDMSLNDIDTQQAATGAGT-----HELLRGNF--NSSNIMKTCLNLAHLRVERL 257
QY 73 -----TTEPDHHTIRLSDET-----NGLVNGAFISHATL 105
DB 258 ILGEFKDKNLEIFESIMEGLCTVIIDFELTYINDFSGLIVKFRCLAN---VSMSL 313
QY 106 --VAFYAWDPVPSLIS-SDIN* 127
DB 314 AGVSKYLEDPVKFKWQSLSTIRC 348

RESULT 8
QYQ0K6 PRELIMINARY; PRI: 835 AA.
AC QYQ0K6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 16, Last annotation update)
DE TOLL-LIKE RECEPTOR 4.
GN TLR4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN-129;
RA Poltorak A., Smirnova I., Chan E.K.L., Bentler B.;
RT "Genetic variation at the TLR4 locus.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN 12;
RP SEQUENCE FROM N.A.
RC STRAIN-129;
RA Poltorak A., Smirnova I., Chan E.K.L., Bentler B.;
RT "Genetic variation at the TLR4 locus.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

QY 18 KAALADDKTNPDTDFMTVTGTTGLTGKGFVEVNLN-NEYAQ-----SNVKSVDG 70
DB 169 KCAALAGGEPVYGVLDVQVRVAG-----NEHISVRLNITYVACMTFGSKNNVNS-HM 220
QY 71 LHTTE--HLDPHLIRLSNCTNNLVNNGAFISHATLVAFKYAMVVPSESTLSSH 125
DB 221 THQLSFGRKYPGTHNPLDQINPIL-----HQLSGIFKYYTIVPQYKSKNW 269

RESULT 10
QYQ0K6 PRELIMINARY; PRI: 423 AA.
AC QYQ0K6;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE SPHINGOMYELIN PHOSPHOLIPID ESTERASE 2 (LIP 3.1.4.12) (NEUTRAL
DE SPHINGOMYELINASE) (NSMASE).
KW Receptor.

```

```

SQ  SEQUENCE 845 AA: 95558 MW: 969485949A220C17 CRC64;

Query Match
Best Local Similarity 11.48; Score 77.5; DB 11; Length 845;
Matches 39; Conservative 14; Mismatches 43; Indels 49; Gaps 8;

QY 21 LAUGKTCNTDPMVTGTTTGLTGKGFVEVNLNNLYAGSNVNVSG-----HGLH---- 72
DB 205 LSLDMSLNDIDTQQAATGAGT-----HELLRGNF--NSSNIMKTCLNLAHLRVERL 257
QY 73 -----TTEPDHHTIRLSDET-----NGLVNGAFISHATL 105
DB 258 ILGEFKDKNLEIFESIMEGLCTVIIDFELTYINDFSGLIVKFRCLAN---VSMSL 313
QY 106 --VAFYAWDPVPSLIS-SDIN* 127
DB 314 AGVSKYLEDPVKFKWQSLSTIRC 348

RESULT 9
QYQ0K6 PRELIMINARY; PRI: 333 AA.
AC QYQ0K6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE BIAAF34232.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Eudicotyledons; Core eudicots; Rosidae; Eucosids 11;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID:3702;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
EX MEDLINE 29277480; PubMed 10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-145 (2000).
DR EMBL; AB022215; BAB17274.1;
SQ  SEQUENCE 333 AA: 37450 MW: 86621350D0B40B9 CRC64;

Query Match
Best Local Similarity 30.58; Score 76.5; DB 10; Length 333;
Matches 36; Conservative 11; Mismatches 44; Indels 27; Gaps 8;

QY 18 KAALADDKTNPDTDFMTVTGTTGLTGKGFVEVNLN-NEYAQ-----SNVKSVDG 70
DB 169 KCAALAGGEPVYGVLDVQVRVAG-----NEHISVRLNITYVACMTFGSKNNVNS-HM 220
QY 71 LHTTE--HLDPHLIRLSNCTNNLVNNGAFISHATLVAFKYAMVVPSESTLSSH 125
DB 221 THQLSFGRKYPGTHNPLDQINPIL-----HQLSGIFKYYTIVPQYKSKNW 269

RESULT 10
QYQ0K6 PRELIMINARY; PRI: 423 AA.
AC QYQ0K6;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE SPHINGOMYELIN PHOSPHOLIPID ESTERASE 2 (LIP 3.1.4.12) (NEUTRAL
DE SPHINGOMYELINASE) (NSMASE).
KW Receptor.

```



```

GN SMPD2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi.
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606.
EN [1]
RE SEQUENCED FROM N.A., CHAPMAN, IFFLAND, AND TISSUE SPECIFICITY
RX MEDLINE:98188255; PubMed:9520418;
PA Tomlik S., Hofmann K., Nix M., Zumbansen M., Stoffel W.;
PT "cloned mammalian neutral sphingomyelinase: functions in sphingolipid
RT signalling?";
KL Proc Natl Acad Sci U S A 95:3638-3643(1998)
CC -1- FUNCTION: CONVERTS SPHINGOMYELIN TO CERAMIDE.
CC -2- CATALYTIC ACTIVITY: SPHINGOMYELIN + H(2O) = N-ACETHYLSPHING-
CC CHOLINE + PHOSPHATE.
CC -3- COFACTOR: MAGNESIUM
CC -4- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -5- MISCELLANEOUS: THIS PROTEIN HAS AN optimum pH of 5.5-6.0.
CC -6- MISCELLANEOUS: THERE ARE TWO TYPES OF SPHINGOMYELINASES: ASM
CC (ACID), AND NSM (NEUTRAL).
CC -7- SIMILARITY: TO OTHER SPHINGOLIPID
CC ENBL: AJ222801; CAA10995.1; ...
DR MM: 603498; ...
KW Hydrolase, Transmembrane, Magnesium.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 354 374 POTENTIAL.
FT METAL 49 49 MAGNESIUM (BY SIMILARITY).
FT SITE 180 180 IMPORTANT FOR SUBSTRATE RECOGNITION (BY
FT SIMILARITY).
FT ACT_SITE 272 272 GENERAL BASE (BY SIMILARITY).
SQ SEQUENCE 423 AA; 4759; MW: 55,523; pI: 5.471; CYS471
Query Match 11.1%; Score 76; DB 4; Length 423;
Best Local Similarity 32.4%; Pred. No. 2.4;
Matches 22; Conservative 4; Mismatches 28; Indels 14; Gaps 1.
QY 49 FEVNI INKVCAGSNVKSVDGLHTTETTHPTFTLSQTNHIVRECAVTEAAVAV 109
I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
DB 233 FGVRIDVLKAVSGFVICKSEFTTGTGDPHSGPLSD-----HEALMAT 278
QY 109 KYANDVPP 116
DB 279 LFRHSPP 286
RESULT 11
ID Q902F5 PPFLIMINARY; PPT: 835 AA.
AC Q902F5 (FEB-1998) (FEB-1998) (FEB-1998) (FEB-1998) (FEB-1998)
DT 01-MAY-2000 (TREMURel 13, Created)
DT 01-MAY-2000 (TREMURel 13, Last sequence update)
DE 01-MAR-2001 (TREMURel 16, Last annotation update)
DE TOLL-LIKE RECEPTOR 4
GN TLR4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi.
OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090.
EN [1]
RE SEQUENCED FROM N.A.
RX MEDLINE:26011445; PubMed:10518109,
PA Fedorukh D.M., Golinsky A., Hajjar A.M., Stevens A., Wilson C.R.,
RA Bassetti M., Adjem A.;
PT "The Toll-like receptor 2 is recruited to macrophage phagosomes and
RT discriminates between pathogens.";
KL Nature 401:811-815(1999).
DR ENBL: AF185285; AF042781; ...
DR InterPro: IPR000157; ...
DR InterPro: IPR000483; ...
DR InterPro: IPR000887; ...
DR InterPro: IPR001611; ...
DR Pfam: PF00560; LRR: 7.

```

```

DR Pfam: PF01463; LRRCT: 1.
DR Pfam: PF01582; TIR: 1.
DE PRINTS: PF00013; IUPUIRPT.
DE PROSITE: PS00159; ALLOLASE_KING_KB2_1; UNKNOWN_1.
DE SMART: SM00255; TIP: 1
KW Receptor.
SQ SEQUENCE 445 AA; 4540; MW: 50,466; pI: 4.4; AF0425; CYS44; CYS45
Query Match 11.0%; Score 75.5; DB 11; Length 835;
Best Local Similarity 26.9%; Pred. No. 6.1;
Matches 39; Conservative 14; Mismatches 43; Indels 49; Gaps 8;
QY 21 TADPTFTTTPVWVTITTALETGTFCEFEVRLRNLYCAGSNVYSC----DGLH 72
I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
DB 205 TSLDTISNPDTIQVAFQGIKI-----HEPLTPGNGF--NSNIMKTKTQNIAGLHRL 257
QY 73 LADLEPHLHLESG; ... RRLVYNGAPISHATL 105
DB 258 ILGEKDERNLETFEPLMEGLADVILDEFKLIYNDSDDIVKHCLEAN----VSAMSL 313
QY 106 --VAFKYANDVTFSEGLIS SDING 127
I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
DB 314 AGVSIKYLEVDVPRKFKWQSLIIRC 338
RESULT 12
ID 042861 PRELIMINARY; PPT: 922 AA.
AC 042861 (FEB-1998) (FEB-1998) (FEB-1998) (FEB-1998) (FEB-1998)
DT 01-JUN-1998 (TREMURel 96, Created)
DT 01-JUN-1998 (TREMURel 96, Last sequence update)
DT 01-MAR-2001 (TREMURel 16, Last annotation update)
DE PUTATIVE HELICASE.
GN SFAC25A8_01C.
OS Schizosaccharomyces pombe (Pission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC S-hii-saccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896.
EN [1]
RE SEQUENCED FROM N.A.
RX STRAIN=972H-;
RA Skellton J., Churcher C.M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCED FROM N.A.
RC STRAIN=972H-;
PA Barrell R.G., Rajandream M.A., Wood V.;
EL Submitter's (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN
DR PMH: AL021809; CAA14951.1; ...
DR InterPro: IPR000330; ...
DR InterPro: IPR001650; ...
DR Pfam: PF00176; SNF2_N; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR SMART: SM00490; HELIC; 1.
KW ATP Binding; Helicase.
SQ SEQUENCE 922 AA; 10450; MW: 90,268; pI: 2.0; CYS45; CYS46; CYS47
Query Match 10.9%; Score 75; DB 3; Length 922;
Best Local Similarity 30.4%; Pred. No. 7.8;
Matches 31; Conservative 9; Mismatches 32; Indels 30; Gaps 4;
QY 45 GKQDFEVLNLYCAGSNVKSVDGLHTTEPL-DEHILPLSDGNNIVNNG----- 97
I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
DB 4 CKPLHIAHGHYATSNVK-----PKPIFPPTANISFATKANKVQGMNSPL 54
QY 98 -----APISHATLVAFKYANDVPPSFISIDIN 126
I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
DB 55 QLLSEMSKRVQATAPIS--SLEHFKQLSDISPSFTSSANSIN 94

```

Search completed: July 18, 2001, 17:28:45
Job time: 77 sec